Models and Methods in Cultural and Social Evolution

Author:
Elliot Aguilar

Supervisor:
Dr. David Lahti

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in the
Ecology, Evolutionary Biology, and Behavior Biology Department

March 2016
Declaration of Authorship

I, Elliot Aguilar, declare that this thesis titled, ‘Models and Methods in Cultural and Social Evolution’ and the work presented in it are my own. I confirm that:

■ This work was done wholly or mainly while in candidature for a research degree at this University.

■ Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.

■ Where I have consulted the published work of others, this is always clearly attributed.

■ Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.

■ I have acknowledged all main sources of help.

■ Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:
“It makes one’s head heavy and giddy, as if one were not looking back down on the receding perspectives of time but rather down on the earth from a great height, from one of those towers whose tops are lost to view in the clouds.”

W.G. Sebald
Abstract

Ecology, Evolutionary Biology, and Behavior

Biology Department

Doctor of Philosophy

Models and Methods in Cultural and Social Evolution

by Elliot Aguilal
Chapter 1

The mathematical study of genealogies has yielded important insights in population biology, such as the ability to estimate the time to the most recent common ancestor (MRCA) of a sample of genetic sequences or of a group of individuals. Here we introduce a model of cultural genealogies that is a step toward answering similar questions for cultural traits. In our model individuals can inherit from a variable, potentially large number of ancestors, rather than from a fixed, small number of ancestors (one or two) as is typical of genetic evolution. We first show that, given a sample of individuals, a cultural common ancestor does not necessarily exist. We then introduce a related concept: the most recent unique ancestor (MRUA), i.e., the most recent single individual who is the earliest cultural ancestor of the sample. We show that, under neutral evolution, the time to the MRUA can be staggeringly larger than the time to MRCA in a single ancestor model, except when the average number of learning opportunities per individuals is small. Our results point out that the properties of cultural genealogies may be very different from those of genetic genealogies, with potential implications for reconstructing the histories of cultural traits.

Chapter 2

A specific goal of the field of cultural evolution is to understand how processes of transmission at the individual level lead to population wide patterns of cultural diversity and change. Previous models of cultural copying and innovation have assumed that traits are independent of one another and essentially exchangeable. But culture has an architecture: traits bear relationships to one another that affect the transmission process itself. Here we introduce an agent based simulation model to explore the effect of cultural architecture on the process of copying and
We construct a space of all possible traits and assign them pairwise, symmetric relationships of compatibility or incompatibility. We then implement different ways for agents to parse these relationships, called filters. We find that introducing this simple architecture leads to novel results. When individuals copy based on a trait’s features (here, its compatibility relationships) they produce smaller, more homogenous cultures than when they copy based on the cultural model. We also find that the average compatibility of a culture produced by some filters is determined by the variance in compatibility in the space of all possible traits, a cultural analog to Fisher’s Fundamental Theorem of Natural Selection. We discuss the implications of considering cultural architecture in the dynamics of cultural change.

**Chapter 3**

Language shift is when a group of speakers adopts a new language, and occurs as part of the larger phenomenon of language contact. It has long been observed that language change is accelerated during shift situations. The standard explanation for this accelerated change has been the introduction of novel linguistic forms by new speakers during the second language acquisition (SLA) process. This hypothesis is based on historical reconstructions of contact situations and has never been formally tested on empirical data. In this paper, we construct an agent-based model to formalize the hypothesis that L2 speakers are responsible for accelerated language change during language shift. In our simulations, a population experiences demographic change via the birth of native (L1) speakers, recruitment of L2 speakers, and death. However, only L2 speakers have the potential to ‘mutate’ (introduce a new variant) on entering the population. We then parameterize the model using demographic data from Maputo, Mozambique—where
rapid shift from Bantu-languages to Portuguese has been occurring for the past forty years—and compare our model predictions to a rare diachronic data set on two linguistic features of Portuguese in Moputu. We find that our basic model is a poor fit to either data set. Next, we modify the model by allowing L2 speakers to introduce a novel variant at any point during their first five years in the population, a feature that represents the fact that the SLA process is not instantaneous (we find support for the five year duration in the literature). We find that the extended SLA model is a good fit to one of our datasets—we discuss plausible reasons for why the other data set is such a mismatch. Finally, we discuss typological differences between contact-induced and non-contact-induced language change and suggest that multiple introductions of a new linguistic variant by different individuals may be the mechanism by which SLA accelerates language change.

Chapter 4

From breeding to flocking, synchrony is an important feature of many social behaviors. Different measures of synchrony have been proposed for different behaviors. Here we test how well some of these measures behave when applied to what we term timing and duration variable scalar behaviors (TDVS). These behaviors are those that may vary in timing (onset) and duration, and which can be characterized by a scalar variable at any time. Using agent-based simulations, we compare the effectiveness of four synchrony measures and show that two, the Dispersion Index of Mean Behavior, and the Kappa coefficient, perform best. We also show that population size affects the statistical interpretation of these measures. Finally, we also include additional results to show the relationship between a number of these measures.
Acknowledgements

The work that appears in the following pages would not have been possible if not for contributions large and small from a number of mentors and collaborators. First, I’d like to thank my co-authors, in order of appearance: Chapter 1, Dr. Stefano Ghirlanda; Chapter 2, Dr. Magnus Enquist and Dr. Alberto Acerbi; Chapter 3, Dr. Anna Jon-And; and Chapter 4, Dr. David Lahti, my advisor. It has been a pleasure working with each of these scientists, and I have benefitted both personally and intellectually from our collaborations.

In addition to fruitful work relationships, I’d also like to thank those who have provided me with guidance, support, and the assurance that I so often found myself in need of during my time as a PhD student. First, I thank my advisor, whose generosity toward me was only exceeded by his intellectual openness. I cannot imagine having reached this point without having an advisor who was so willing to let me follow my curiosity to the utmost. After all these years I’ve only scratched the surface of his understanding of evolution and so many other things.

I would like to thank Dr. Stefano Ghirlanda, who has been a friend, a mentor, and a guide through the tortured but gratifying work of trying understanding nature through equations. While he appears as a co-author on the first chapter, his influence and advice is woven through all the rest.

Dr. Mike Hickerson deserves special distinction for the enormous support and advice he has provided me, and so many other students, over the past years. I have run up a debt of gratitude so large I can only hope to repay it by following his example in the future.

I would also like to thank my patient committee members, Dr. Ofer Tchernokovski and Dr. Nina Fefferman, for their attention and for agreeing to review the following papers.
Finally, I would like to thank my family: my Mother and Father, without whose love and efforts the dream of becoming a scientist would have remained just that; my grandparents for so many trips to the library and the American Museum of Natural History—surely those trips are the start of all the work that appears here; my brother for always giving me someone to chase; Giovanna Calvino for my best collaboration by far; and finally, my daughter, Violette, who constantly reminds me that I should be asking ‘why’.
Contents

Declaration of Authorship i

Abstract iii

Acknowledgements vii

Contents ix

List of Figures xii

List of Tables xiv

Abbreviations xvii

1 Modeling the genealogy of a cultural trait 1
  1.1 Introduction ...................................................... 2
  1.2 Model ............................................................... 4
    1.2.1 Defining the MRUA concept ............................... 4
    1.2.2 Learning Model .............................................. 6
    1.2.3 Genealogical Model ........................................ 7
  1.3 Results .............................................................. 10
    1.3.1 Model Analysis ............................................... 10
    1.3.2 Time to the MRUA ........................................... 11
    1.3.3 Model comparison ........................................... 13
  1.4 Discussion ......................................................... 15

2 Evolution of Cultural Systems 22
  2.1 Introduction ...................................................... 24
  2.2 Model description ............................................... 29
    2.2.1 Trait Universe ............................................... 29
    2.2.2 Filters ....................................................... 30
    2.2.3 Cultural Copying Model .................................... 32
2.3 Results

2.3.0.1 How does the underlying architecture of relationships between traits affect the size and diversity of culture? 34

2.3.0.2 How do different modes of individual transmission affect the evolution of culture at a system level? 36

2.3.1 What causes cultural systems to diverge? 39

2.4 Discussion 40

2.4.1 Filters 40

2.4.2 Model assumptions and future directions 42

3 An evolutionary model of contact induced language change

3.1 Introduction 48

3.2 Methods 50

3.2.1 Model of language change 50

3.2.2 Demographic change/growth 52

3.2.2.1 Maputo language change 53

3.3 Results 55

3.4 Discussion 57

4 On measuring behavioral synchrony

4.1 Introduction 65

4.1.1 Synchrony Measures 66

4.1.1.1 Kempenaer Index 67

4.1.1.2 Simpson Index 67

4.1.1.3 Kappa Coefficient 68

4.1.1.4 Dispersion Index of mean behavior/Average pairwise covariance 69

4.2 Methods 70

4.3 Results 74

4.3.1 Simple Models 74

4.3.2 Complex Models 76

4.3.3 Synchrony-Inducing conditions 77

4.4 Discussion 77

A Expected time to MRUA

A.1 Variance in time to MRUA 81

B Expected time to MRUA in Strimling et al. model 83

C Algorithm for simulating the genealogical process 85

D System Measures and Cultural Systems Simulation Model details

D.0.1 Compatibility calculations 87

D.0.1.1 Internal Compatibility 88

D.0.1.2 Between individual compatibility 88
List of Figures

1.1 Sample learning genealogy for a cultural trait (e.g. a folksong). While B, C, D, and E are all ancestral to A, only B and C share a common ancestor in E. D has no learning lineages that intersect with the other individuals in the genealogy. . . 5

1.2 Example of a cultural genealogy from a simulation of the model in the text, with \( a = 1.5, r = 1 \). Horizontal lines indicate individuals through time, with solid lines highlighting individuals within the sample being tracked. Arrows indicate learning events and \( \times \)'s indicate replacement. Times, \( t_j \), represent the occurrence of the \( j \)th event (birth or replacement), shown here at equal intervals for convenience. Five individuals are followed through 15 successive events. The initial sample has been arbitrarily defined, for illustration, as individuals 3 and 5 at \( t = 0 \). Starting from the right side, the graph is considered backward in time. Any individual located at the base of an arrow incident on a sample member is added to the sample. Sample members who are replaced are removed from the sample. Individual \( i \) at time \( t_j \) is an ancestor of individual \( k \) at time \( t_l \) if it is possible to reach point \((k, t_l)\) from \((i, t_j)\) following forward arrows. The open circle indicates the most recent unique ancestor of the sample (MRUA, see text). Sample size is tracked in the plot below the genealogy. . . . . . . . . . . . . . . 18

1.3 Expected times to MRUA in samples of 2 and \( N \) individuals in populations of size \( N \). Parameters values: \( a = 1.5 \) (left), \( a = 2 \) (center), \( a = 5 \) (right). Time is measure in units of expected lifetime (\( r = 1 \)), see section 1.2.2. The vertical scale is the same to facilitate comparison between the graphs. . . . . . . . . . . . . . . 18

1.4 Ratios of expected MRUA times, \( E\tau_N/E\tau_2 \), for \( N = 2, 10, 50, 100 \). In all cases, as \( a \) increases, the ratio approaches 1. . . . . . . . . . . . . . . . . . . . . . . . . . . 19

1.5 Change of sample size with time in model simulations (mean: dotted lines; standard deviation: gray area) and in the deterministic approximation (solid lines, equation 1.9) for three values of learning rate parameter \( a \in \{1.5, 2, 5\} \) and two values of initial sample size, \( X_0 = 2 \) (above), \( X_0 = 100 \) (below), for a population of \( N = 100 \) individuals. Replacement rate is \( r = 1 \). All simulations that reach MRUA in the plotted time window are shown in light gray. One-thousand simulations were performed for each choice of parameters. . . . . . . . . . . . . . 20

1.6 Numerical determination of the approximate relationship in (1.11). The left panel shows that \( E\tau_N \) is approximately exponential in \( k^* \), as defined in (1.10), \( E\tau_n \approx 10^{bk^*} \). The right panel shows that \( b \) in the last expression is a logarithmic function of the learning rate parameter \( a \). Numerical fits have been obtained with the \texttt{lml} function of the \texttt{R} software, version 3.0.0, applied to log-transformed values as appropriate [83]. . . . . . . . . . . . . . . . . 20
1.7 Times to MRUA in the multiple ancestry model (solid lines) and single ancestor model (dashed line) for an initial sample size of 2. .................. 21

2.1 (A) Average culture size (across ten runs) with increasing $c$ in the trait universe. (B) Average repertoire size (across ten runs) with increasing $c$ in trait universe. . 44

2.2 Average compatibility of cultural systems (A) within individuals, (B) between individuals, (C) population-wide. . . . . . . . . . . . . . . . . . . . . . 45

2.3 Filter-Neutral compatibility vs. coefficient of variation in number of compatible trait relationships. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 46

2.4 Nearest neighbor distribution under TF and MF in universe with $c=.1$ . . . . 46

3.1 Change in the population mean usage of the mutant variant (averaged across ten runs) over 100 years. The other model parameter values are birth rate, $b=.007$, recruitment rate, $r=.063$, learning parameter, $l=.1$. For fixed population runs, $N=100$, while for expanding populations, $N_0 = 100$, with a growth rate of $g=b+r-d=.05$; $\mu$ is the mutation rate among L2 speakers. The trajectories of the expanding population means show more stability due to lower variance across runs. .................. 61

3.2 Comparison of verb and preposition usage data with the mean model trajectories that most closely approached the data points for 1997, $(l=.05, \mu=.05),(l=.5, \mu=.05),(l=.7, \mu=.05),(l=.9, \mu=.05)$. Both data sets show greater spread than predicted by the model in the third growth phase (1997-2007), though the slope of the preposition data is reduced in qualitative agreement with the model. The verb data actually shows an increase in the spread of the novel variant, which conflicts with the model predictions. .................. 62

3.3 Comparison of verb and preposition usage data with the model 2 mean trajectories that most closely matched the data points. The preposition use data fits the model predictions for $\mu = .0078$. The verb form data for 1997 fell within the 95% confidence intervals for the $\mu = .0063$ runs, but outside them for 2007. .................. 63

4.1 Fitted null distributions for the Kappa coefficient for complex conformity bias runs with increasing population size. .................. 79
### 3.1 Demographic data on Portuguese speakers in Maputo, Mozambique.  

#### 4.1 Correlation of success rate with $w_e$ in simple models.  

#### 4.2 Correlations between success rate and $w_e$ for all parameter pairs in the simple simulations, where success rate is defined as the proportion of times that a measure detected synchrony with $p > .05$. $\zeta^*$ and $\eta^*$ represent $NA$ values due to the success rate being 0 and 1, respectively.  

#### 4.3 Correlation of success rate with $w_e$ in complex models.  

#### 4.4 Correlations between success rate and $w_e$ for all parameter pairs in the complex simulations, where success rate is defined as the proportion of times that a measure detected synchrony with $p > .05$. $\zeta^*$ and $\eta^*$ represent $NA$ values due to the success rate being 0 and 1, respectively.  

#### F.1 Estimated rate of recruitment of L2 speakers for Maputo Portuguese-speaking population.  

#### J.1 Dispersion Index of Polarization, $N=10$  

#### J.2 The proportion of runs for which the Dispersion Index of Polarization detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 10.  

#### J.3 Simpson Index, $N=10$  

#### J.4 The proportion of runs for which the Simpson Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 10.  

#### J.5 Kempenaer Index, $N=10$  

#### J.6 The proportion of runs for which the Kempenaer Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 10.  

#### J.7 Kappa coefficient, $N=10$  

#### J.8 The proportion of runs for which the Kappa coefficient detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 10.  

#### J.9 Dispersion Index of Polarization, $N=50$
J.10 The proportion of runs for which the Dispersion Index of Polarization detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 50. 115

J.11 Simpson Index, N=50 116

J.12 The proportion of runs for which the Simpson Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 50. 116

J.13 Kempenaer Index, N=50 117

J.14 The proportion of runs for which the Simpson Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 50. 117

J.15 Kappa coefficient, N=50 118

J.16 The proportion of runs for which the Kappa coefficient detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 50. 118

J.17 Dispersion Index of Polarization, N=100 119

J.18 The proportion of runs for which the DIP detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100. 119

J.19 Simpson Index, N=100 120

J.20 The proportion of runs for which the Simpson Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100. 120

J.21 Kempenaer Index, N=100 121

J.22 The proportion of runs for which the Kempenaer Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100. 121

J.23 Kappa coefficient, N=100 122

J.24 The proportion of runs for which the Kappa coefficient detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100. 122

K.1 Dispersion Index of polarization, complex model, N=10 124

K.2 The proportion of runs for which the Dispersion Index of Polarization detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100. 124

K.3 Simpson Index, complex model, N=10 125

K.4 The proportion of runs for which the Simpson Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100. 125

K.5 Kempenaer Index, complex model, N=10 126

K.6 The proportion of runs for which the Kempenaer Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100. 126
K.7  Kappa coefficient, complex model, N=10 ................................. 127
K.8  The proportion of runs for which the Kappa coefficient detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100. .................. 127
K.9  Dispersion Index of polarization, complex model, N=50 ...................... 128
K.10 The proportion of runs for which the Dispersion Index of Polarization detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 50. .......................... 128
K.11 Simpson Index, complex model, N=50 ...................................... 129
K.12 The proportion of runs for which the Simpson Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 50. .......................... 129
K.13 Kempenaer Index, complex model, N=50 ................................... 130
K.14 The proportion of runs for which the Kempenaer Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 50. .......................... 130
K.15 Kappa coefficient, complex model, N=50 .................................. 131
K.16 The proportion of runs for which the Kappa coefficient detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 50. .......................... 131
K.17 Dispersion Index of Polarization, complex model, N=100 ................ 132
K.18 The proportion of runs for which the Dispersion Index of Polarization detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100. .......................... 132
K.19 Simpson Index, complex model, N=100 ..................................... 133
K.20 The proportion of runs for which the Simpson Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100. .......................... 133
K.21 Kempenaer Index, complex model, N=100 ................................ 134
K.22 The proportion of runs for which the Kempenaer Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100. .......................... 134
K.23 Kappa coefficient, complex model, N=100 ................................ 135
K.24 The proportion of runs for which the Kappa coefficient detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100. .......................... 135
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRCA</td>
<td>Most Recent Common Ancestor</td>
</tr>
<tr>
<td>MRUA</td>
<td>Most Recent Unique Ancestor</td>
</tr>
<tr>
<td>SLA</td>
<td>Second Language Acquisition</td>
</tr>
<tr>
<td>TDVS</td>
<td>Time and Duration Variable Scalar (behavior)</td>
</tr>
<tr>
<td>DIMB</td>
<td>Dispersion Index of Mean Behavior</td>
</tr>
<tr>
<td>APC</td>
<td>Average Pairwise Covariance</td>
</tr>
</tbody>
</table>
For my daughter.
Chapter 1

Modeling the genealogy of a cultural trait

Abstract

The mathematical study of genealogies has yielded important insights in population biology, such as the ability to estimate the time to the most recent common ancestor (MRCA) of a sample of genetic sequences or of a group of individuals. Here we introduce a model of cultural genealogies that is a step toward answering similar questions for cultural traits. In our model individuals can inherit from a variable, potentially large number of ancestors, rather than from a fixed, small number of ancestors (one or two) as is typical of genetic evolution. We first show that, given a sample of individuals, a cultural common ancestor does not necessarily exist. We then introduce a related concept: the most recent unique ancestor (MRUA), i.e., the most recent single individual who is the earliest cultural ancestor of the sample. We show that, under neutral
evolution, the time to the MRUA can be staggeringly larger than the time to MRCA in a single ancestor model, except when the average number of learning opportunities per individuals is small. Our results point out that the properties of cultural genealogies may be very different from those of genetic genealogies, with potential implications for reconstructing the histories of cultural traits.

1.1 Introduction

One of the most intriguing questions in the study of cultural evolution is to what extent cultural data can inform us about the past. Researchers have used cultural data sets to estimate the dates of important historical and evolutionary events, such as the age of human language [65], the spread of ethno-linguistic groups [34], and the origin of the capacities for cultural transmission [58]. In particular, phylogenetic methods from molecular systematics have been applied to cultural datasets to infer the ages of cultural “taxa,” such as languages, and the relationships between them [33, 54, 72, 85].

These methods are a useful complement to archaeological and historical evidence of cultural change, but rely on data from assemblages of many cultural traits typified at a population level, and thus investigate macroevolutionary cultural change. For example, these studies usually treat languages as characteristic sets of lexical, phonemic, and grammatical features shared by all members of a population. Individual cultural traits—e.g. words, techniques for making hand axes, oral literature, folksongs, childrearing practices, etc.—may have unique histories that are lost when considering only the larger systems they constitute at the macroevolutionary level. While we acknowledge the difficulty inherent in defining a ‘single’ cultural trait—indeed,
many traits can be thought of as combinations or systems of component traits—we assert that whatever the definition, single traits spread due to the microevolutionary process of individual to individual transmission. In this paper, we develop a model of the history of a single cultural trait transmitted between individuals based on population size, mode of trait transmission, and current prevalence of the trait.

In order to investigate the history of a cultural trait we require a model of the genealogical process for learned traits. In population biology, the mathematical study of genealogies has flourished since the introduction of Kingman’s “coalescent,” a retrospective model of how lineages of gene copies merge in common ancestors [50–52]. The coalescent predicts the statistical properties of genealogies and thereby enables the estimation of other aspects of population history such as ancestral population size. A central concept in genetic genealogies is the most recent common ancestor (MRCA), i.e., the most recent individual in the past whose gene copy is ancestral to all those in the present. The coalescent estimates that, for a population of size $N$, the expected time to MRCA is of the order of $N$ generations [46]. Might retrospective models from population genetics shed light on the behavior of cultural trait genealogies, just as molecular systematics has done for cultural taxa? We argue that fundamental differences in the nature of cultural transmission make genetic genealogical models unsuited for culture. While genes are inherited uniparentally (or, in the case of recombination, biparentally), cultural traits can be inherited from multiple ancestors. For example, an oral story may be learned in repeated episodes from a number of sources.

Here we introduce a simple, neutral model of cultural genealogies in order to investigate the relationship between the individual-to-individual nature of cultural transmission and historical inference based on cultural data. Specifically, we model the history of learning events pertaining
to a single cultural trait possessed by a group of individuals. Each individual has a variable, potentially large number of cultural ancestors—in contrast to models of genetic genealogies in which each individual has one or two ancestors. Individuals form their traits based on information inherited from their multiple ancestors—thus, a learning event may represent the partial or complete transmission of a trait. We first show that, for a sample of individuals drawn from a population, a cultural MRCA does not always exists. That is, the genealogical lines of the sampled individuals may fall into two or more disjoint sets, without ever intersecting. Alternately, we define a concept related to the MRCA but that identifies an individual that is guaranteed to have existed. We call this individual the most recent unique ancestor (MRUA), defined as the most recent single individual who is a source of cultural information for individuals in the sample. For any random sample of a population of fixed size and undergoing neutral cultural evolution, we show that the expected time to the MRUA grows hyper-exponentially with population size, and thus, above certain parameter values, is much larger than the time to MRCA in a uniparental model. This result highlights the need for more theory on historical inference based on cultural data.

1.2 Model

1.2.1 Defining the MRUA concept

The distinction between MRUA and MRCA exists because of differences between cultural and genetic inheritance. A single gene copy is inherited uniparentally, and has its origin in a single individual. This means that as we move backwards in time lineages of copies of a gene will eventually converge in the common ancestor. Even in the event of recombination, the gene copy
has at most two parents and those lineages too will eventually merge in a common ancestor. By contrast, a person’s cultural trait may have inputs from many sources. The original trait could have arisen in a single individual or among a group of individuals. In the latter case it is clear that no single MRCA exists. However, even if a trait has a single origin it may still be modified by subsequent inputs of information. The individuals who contribute this additional information may be “dead ends” in the genealogy, since their own learning lineages may never merge with the others in the sample. Figure 1 depicts an example of a cultural genealogy as a directed graph. The nodes represent individuals in the genealogy, and arrows show ancestor-descendant relationships (arrows point from ancestor to descendant). At the top of the figure we have individual A, who possesses a cultural trait, let us say knowledge of a specific folksong. At some point in the past, A learned her song from B, C, and D. Prior to that, B and C learned it from E, thus E is their common ancestor. However, Individual D served as a learning model for A—perhaps contributing a new verse—without having learned anything from E. Thus, E is not a common ancestor for everyone in the genealogy, but she is still the earliest single individual who contributes to the song learned by A. For this reason, we refer to E as the most recent unique ancestor, or MRUA.

In genetic evolution, origin and common ancestry coincide. As we have seen, in cultural evolution
there may be no common ancestor, since: (1) individuals may independently invent a trait that already exists (e.g. Leibniz’s and Newton’s independent inventions of the calculus), and go on to serve as learning models; and (2) individuals may independently invent some information that gets incorporated into a trait even without possessing it themselves, as in the example of D above. Therefore, the origin of a trait for some group of individuals is represented by the MRUA, which plays a role analogous to the that of the MRCA in the neutral coalescent.

1.2.2 Learning Model

To investigate the properties of cultural genealogies, we begin with a simple model of cultural transmission. We consider a fixed population of $N$ individuals that evolves according to a continuous-time Markov process in which only two events can occur:

**Replacement:** A randomly selected individual is removed from the population and is replaced by a naive individual.

**Learning:** A randomly selected individual learns from another randomly selected individual.

The time between successive replacement events is exponentially distributed with rate $r$ per individual, and the time between successive learning events is exponentially distributed with rate $a$ per individual. Equivalently, we can say that replacement and learning events occur according to a Poisson process with rate $(a+r)N$. A given event is a replacement with probability $r/(a+r)$, and a learning event with probability $a/(a+r)$. The expected lifetime of individuals is $1/r$, and the expected number of learning events per lifetime is $a/r$.

This model allows individuals to acquire multiple cultural ancestors via multiple learning events in the course of a lifetime. We define a learning event as the transmission of some amount of
information about a trait from one individual to another. We do not assume that the event
causes the exact replication of a cultural trait, and we leave unspecified exactly how an individual
forms her cultural trait from multiple inputs. While cultural transmission can be conceptualized
in many ways [41], here we are concerned only with tracking genealogies of a single trait. In
other words, regardless of the details of how cultural transmission occurs, we can consider
anyone from whom an individual has learned as the individual’s cultural ancestor. Note that
in a multiple-ancestry model it is important to distinguish learning events from ancestors. Any
number of learning events with the same ancestor, in fact, results in a single lineage in the
genealogy (see Section 2.3).

1.2.3 Genealogical Model

The aim of a theory of cultural genealogies is to infer the statistical properties of genealogies from
knowledge of the process that generates them, which, in the present case, is the simple cultural
transmission process introduced above. We are not tracking the spread of a cultural variant, but
modeling the history of learning events that have resulted in a group of individuals possessing
a given trait. We imagine encountering a population in the present with no information about
the history of learning relationships between its members, and we want to know how far back
in time we would expect to find the MRUA of a sample of \( n \) individuals who share a trait in
common.

To calculate the expected time to MRUA we define the random variable \( X_t \) as the number
of ancestry lines existing at a time \( t \) before the present that lead to individuals in the initial
sample. Starting from an initial sample size, \( X_0 = n \), we define the time to MRUA \( , \tau_n \) as the
smallest time for which \( X_t = 1 \):

\[
\tau_n \overset{\text{def}}{=} \inf \{ X_t = 1 \mid X_0 = n \} \tag{1.1}
\]

In our model there are only two events that affect \( X_t \):

1. An individual in the sample learns from an individual outside the sample. This leads to the branching of an ancestry line, increasing sample size by 1. These events occur with rate

\[
\text{learning rate} \times \Pr (\text{learner in sample}) \times \Pr (\text{ancestor outside sample}) = aN \times \frac{X_t}{N} \times \frac{N - X_t}{N - 1} \tag{1.2}
\]

2. An individual in the sample is replaced by a naïve individual. This leads to the interruption of an ancestry line, decreasing sample size by 1. These events occur with rate

\[
\text{replacement rate} \times \Pr (\text{replacement occurs in the sample}) = rN \times \frac{X_t}{N} \tag{1.3}
\]

Other events, such as replacement of an individual outside the sample or learning between individuals who are in the sample, do not affect \( X_t \).

As a consequence of learning and replacement events, sample size changes according to a continuous time birth-death process whose states are in the subset of integers \( \{0, N\} \) (though we are only concerned with the process until it reaches state \( X_t = 1 \)), and with state-dependent
birth and death rates (i.e. $X_t = k$), $\lambda_k$ and $\mu_k$, given by equations (1.2) and (1.3):

$$\lambda_k = ak \frac{N - k}{N - 1} \quad (1.4)$$

$$\mu_k = rk \quad (1.5)$$

The time between events is exponentially distributed with parameter $\kappa_k = \lambda_k + \mu_k$. The time evolution of $X_t$ governed by the rates given above is the genealogical process.

Fig. 1.2 shows the results from a simulation of the genealogical process for a population $N = 5$ (see Appendix D for simulation algorithm). The lower plot shows the time evolution of $X_t$ over a series of events resulting in MRUA. The graph above, inspired by the percolation diagram introduced by Krone & Neuhauser (1997) [55] to represent the coalescent process with selection, shows the simulated genealogy. Horizontal lines represent individuals, and time moves backwards as we move from right to left. Vertical arrows are learning events with the learner at the head of the arrow (i.e. information ‘flows’ in the direction of the arrow), while $\times$-marks indicate replacement of an individual. Beginning at the rightmost end of the figure, we consider an initial sample of individuals 3 and 5 who possess some cultural trait. Solid lines indicate lineages that are part of our sample, or events that affect sample size. Thus, $X_t$ is equal to the number of solid horizontal lines at $t$. Moving left along the lineages and following the arrows in reverse, we arrive at individual 1, the MRUA.

The genealogical process generated by our model bears a superficial resemblance to the dual process introduced by Krone and Neuhauser (1997) and a conceptual similarity to the recombination graph [35, 36]; all three are equivalent to a coalescing, branching process on a complete graph of order $N$. Unlike the result given below, which is explicit for finite population size,
N, Krone & Neuhauser find an expected time expression in the limit $N \to \infty$, using diffusion limit forms of the state-dependent coalescence and branching rates. Similarly, the recombination graph uses diffusion scalings of the transition rates to solve for expected times in infinite populations. There is an important similarity between the recombination graph and our own process, which is that it allows for gene sequences to have upwards of two parents and thus, for finite populations, could be seen as a special case of our multi-ancestor model. However, as will be mentioned below, our process leads to results that differ from those of the recombination graph.

1.3 Results

1.3.1 Model Analysis

Let $X_t = k$; the process will remain at $k$ for an expected time $1/\kappa_k$ before jumping to state $X_{t+s} = k + 1$ with probability $\lambda_k/\kappa_k$, or state $X_{t+s} = k - 1$ with probability $\mu_k/\kappa_k$ ($s$ is the actual waiting time, exponentially distributed with mean $1/\kappa_k$). Therefore, the expected value of $\tau_k$ is given by the following second-order difference equation [62]:

$$E[\tau_k] = \frac{1}{\kappa_k} + \frac{\lambda_k}{\kappa_k} E[\tau_{k+1}] + \frac{\mu_k}{\kappa_k} E[\tau_{k-1}]$$

(1.6)

for $2 \leq k \leq N - 1$, and with boundary conditions

$$E[\tau_1] = 0, \quad E[\tau_N] = \frac{1}{\kappa_N} + E[\tau_{N-1}]$$

(1.7)
Chapter 1. *Modeling the genealogy of a cultural trait*

The first condition reflects the fact that \( n = 1 \) is the state corresponding to the convergence of all sample lineages—excluding those that have terminated in “dead ends”—in the MRUA; the second condition reflects the fact that the maximum value of \( X_t \) is the population size, \( N \). The solution to (1.6) with these boundary conditions is given by (see Appendix A),

\[
E[\tau_n] = \frac{1}{r} \sum_{i=2}^{n} \sum_{j=i}^{N} \frac{1}{j} \left( \frac{a}{r(N - 1)} \right)^{j-i} \frac{N^j \cdots (N - 1)}{N^i} \tag{1.8}
\]

where \( N^j \) is the “descending factorial” or “falling power” \[32\], defined by \( N^j \equiv N(N - 1)\cdots(N - j + 1) \). A similar calculation leads to the variance and higher moments (see Appendix B).

### 1.3.2 Time to the MRUA

Equation (1.8), though exact, is not transparent. In Fig. 1.3 we show numerical calculations of times to MRUA for parameter values intended to represent cultural traits that are learned in only a few learning events (small \( a \)) or from a large number of events (larger \( a \), expected for many human cultural traits). MRUA times grow faster than exponentially with population size, reaching staggering magnitudes even for a relatively modest \( a = 5 \). We can also see that a small increase from an average 1.5 to 2 learning events leads to a \( 10^4 \)-fold increase in MRUA times. These properties stand in contrast with the time to the genetic MRCA, which is of the order of population size, \( N \) for large samples \[50\].

Fig. also suggests that increasing \( a \) decreases the difference in MRUA times between different initial sample sizes. To clarify this result, Fig. 4 plots the ratio \( E[\tau_n]/E[\tau_2] \) for a range of population sizes. As the learning parameter \( a \) increases, the ratio approaches 1, and expected MRUA times
do not differ between initial sample sizes. The learning parameter $a$ is a feature of the trait being modeled; some traits will on average be learned in more or fewer learning episodes. Thus, for a trait that is acquired in many learning episodes (i.e. high value of $a$), we should not expect the MRUA to be any older because the trait is popular than we would if it were very rare.

To shed light on the dependence of MRUA times on population parameters, we approximate the stochastic process with a deterministic process obtained by considering expected changes in sample size. When the stochastic process is in state $k$ the expected change in sample size is

$$E[\Delta k] = \frac{\lambda_k - \mu_k}{\lambda_k + \mu_k} = \frac{a(N-k) - r(N-1)}{a(N-k) + r(N-1)}$$

(1.9)

The deterministic approximation is obtained by using this expected change to recursively update sample size. Figure 1.5 shows that equation (1.9) accurately describes the mean evolution of the stochastic process, obtained from simulation. The deterministic approximation reaches an equilibrium value, $k^*$, which corresponds to the quasi-stationary equilibrium of the stochastic process and is independent of initial sample size, $n$. Setting $E(\Delta k) = 0$ and solving for $k$ we get the equilibrium sample size

$$k^* = N - \frac{r}{a}(N-1)$$

(1.10)

Thus the equilibrium value of the deterministic process decreases linearly with $r/a$. After a transient phase determined by the initial sample size, the stochastic process hovers around its quasi-stationary equilibrium until a fluctuation brings it down to the state $n = 1$. Higher values of $k^*$ correspond to longer times to the MRUA because larger fluctuations are necessary to reach $k = 1$. Numerical analysis indicates that time to the MRUA is an exponential function of $k^*$,
namely (see Fig. 1.6):

\[ E[\tau_n] \approx 10^{0.7k^* \log a} \]  

(1.11)

The above equation does not take into account the variation of \( E\tau_n \) with initial sample size, which is significant only for small learning rates, as mentioned in the previous paragraph (see Fig. 1.3). The deterministic approximation shows more clearly the role of \( a \) and \( r \) in determining time to the MRUA. Increasing \( a \), the expected number of learning events, and thus the potential for multiple ancestry, increases the quasi-stationary equilibrium sample size, prolonging time to MRUA.

### 1.3.3 Model comparison

To further explore the implications of multiple ancestry we compared our model with a uniparental model of cultural transmission. The uniparental model we select is that introduced by Strimling et al. (2009) to investigate the accumulation of cultural traits in a fixed population. While the model was originally described in discrete time, we consider a continuous time version. Ignoring mutation and imperfect copying, the model proceeds in the following steps:

1. A randomly selected individual leaves the population.

2. A new individual replaces her and chooses another member of the population at random to be her sole cultural parent.

The times between events are exponentially distributed with parameter \( r \), with \( 1/r \) being the average individual lifetime; for simplicity we set \( r = 1 \) for both models.
Chapter 1. *Modeling the genealogy of a cultural trait*

The model of Strimling et al. is similar to a Moran model of reproduction, but with an important exception: the individual that leaves the population cannot be chosen as a cultural parent by her replacement. All lineages will eventually converge in an MRCA because, as in genetic evolution, information is inherited uniparentally. Therefore, MRCA and MRUA are equivalent in the Strimling et al. model, and we will only refer to MRUA in the following discussion. In continuous time, with $r = 1$, we arrive at the following expression for the time to MRUA for a sample of $n$ individuals (see Appendix B)

$$E\tau_n = (N - 1) \sum_{j=2}^{n} \frac{1}{j(j - 1)}$$

(1.12)

This result is nearly identical to that of the standard neutral coalescent. Figure 1.7 shows a comparison of MRUA times in the multiple ancestry and Strimling et al. models with increasing $N$ and an initial sample, $X_0 = 2$. In our model, the time to MRUA grows hyper-exponentially with population size, typically, while it grows linearly in the Strimling et al. model. However, when learning opportunities are few ($a$ smaller than a critical value that depends on population size), the expected times are actually longer for the Strimling et al. model than for our own. For example, for $a = 1$, meaning individuals have on average one learning event per lifetime, the multiple ancestry model times are shorter by an order of magnitude for populations larger than 10, for all values of initial sample size up to $N$. Given the general effect of multiple ancestry on extending MRUA times, it is interesting to see that in some cases the multiple ancestry model can actually lead to MRUA times that are shorter than in the single-ancestor model. This stands in contrast to the recombination graph mentioned above, whose minimum expected times are equivalent to the standard neutral coalescent.
Chapter 1. *Modeling the genealogy of a cultural trait*

1.4 Discussion

We constructed what is to our knowledge the first model of cultural transmission to explore the behavior of cultural genealogies. The major advantage of our model is its generality; we tracked the learning lineages without making any assumptions about the nature of transmission (e.g., discrete vs. continuous units) or about how cultural phenotypes emerge from learned information. The model has two key hypotheses: neutrality and multiple inheritance.

Neutrality, i.e., random choice of individuals for replacement and learning independent of their biological and cultural traits, is not realistic in general. Cultural time series often appear to deviate from neutrality [3, 37], though not invariably [11]. Many mechanisms could cause such deviations: some cultural variants may be acquired more easily (or discarded less easily), some individuals may be preferred as cultural models, cultural traits may influence survival and reproduction, and so on [3, 16, 40, 80]. Future research will have to assess how the cultural genealogical process is affected by these factors, and by others such as variable population size (e.g., a population expansion will appear, looking back in time, as a reduction in the number of possible ancestors, which should shorten times to the MRUA.) However, by assuming neutrality we have been able to isolate the effect of the other main hypothesis: multiple ancestry.

Unlike neutrality, multiple ancestry is a fundamental aspect of cultural transmission [16, 20, 27] and thus a necessary ingredient of any model of cultural genealogies. We have shown that multiple ancestry leads to genealogies that behave very differently from those produced by haploid and diploid inheritance. First, the concept of a most recent common ancestor, familiar from population genetic theory, appears less useful in the study of cultural genealogies, because these may lack a common ancestor altogether. We proposed to overcome this difficulty
by studying instead the “unique” ancestor of a sample of individuals—the most recent single
individual who contributed cultural information to at least one of the lineages leading to the
sample. This individual is guaranteed to have existed, and coincides with the most recent
common ancestor when the latter exists. Our model demonstrated that multiple inheritance
generally causes times to MRUA to be much longer than in a uniparental model.

MRUA and MRCA are analogous because they both identify an ancestral individual at the root
of a genealogy. However, other definitions of “common ancestor” are possible. For example, a
common ancestor can be defined as any individual who appears in the genealogies of all those
sampled. This is the lay definition of common ancestry, according to which a grandmother is
a common ancestor of all her grandchildren. Chang (1999) showed that the time to such an
ancestor for a random sample of a large diploid population is of the order of $\log_2 N$ generations
for large $N$, thus much shorter than the time to MRCA in uniparental models, which is of the
order of $N$ generations, and of course shorter than time to MRUA in our model. We leave it
to future work to determine the expected time to “cultural grandmothers”, though we suspect
that the times will be even shorter than the diploid result, due to the branching nature of our
process.

Our model may be applicable to questions of genealogical processes in cultural transmission,
such as oral story transmission. The model incorporates the possibility of learning from multiple
individuals, learning multiple times from the same individual, and independent contribution,
which are all factors in the transmission of oral stories [87]. Estimating the parameters of
the model, “lifetime” of the active storyteller/learner, and rate of learning events, might lead
to estimates of the time of MRUA for a story, which could be compared with independent
historical evidence to validate the model. Experimental setups could also be used to investigate
the time to the MRUA of stories as a function of parameter values, and to compare these times to the case of uniparental transmission. As there is a growing body of experimental work on oral transmission and cultural evolution [9, 28], this is an especially promising direction for empirical application of our model.

In conclusion, our work represents a first step toward a theoretical foundation for historical inference based on cultural data. We investigated the history of an individual cultural trait by explicitly modeling the individual-to-individual transmission process with multiple ancestry. Many other properties of cultural genealogies remain to be explored, such as the effects of biased transmission, and the effect of historical fluctuations in the popularity of a trait. Further work is needed to ascertain the impact of the unique features of cultural inheritance on the theory and practice of reconstructing the histories of cultural traits.
Chapter 1. Modeling the genealogy of a cultural trait

Figure 1.2: Example of a cultural genealogy from a simulation of the model in the text, with $a = 1.5, r = 1$. Horizontal lines indicate individuals through time, with solid lines highlighting individuals within the sample being tracked. Arrows indicate learning events and $\times$'s indicate replacement. Times, $t_j$, represent the occurrence of the $j$th event (birth or replacement), shown here at equal intervals for convenience. Five individuals are followed through 15 successive events. The initial sample has been arbitrarily defined, for illustration, as individuals 3 and 5 at $t = 0$. Starting from the right side, the graph is considered backward in time. Any individual located at the base of an arrow incident on a sample member is added to the sample. Sample members who are replaced are removed from the sample. Individual $i$ at time $t_j$ is an ancestor of individual $k$ at time $t_l$ if it is possible to reach point $(k, t_l)$ from $(i, t_j)$ following forward arrows. The open circle indicates the most recent unique ancestor of the sample (MRUA, see text). Sample size is tracked in the plot below the genealogy.

Figure 1.3: Expected times to MRUA in samples of 2 and $N$ individuals in populations of size $N$. Parameters values: $a = 1.5$ (left), $a = 2$ (center), $a = 5$ (right). Time is measure in units of expected lifetime ($r = 1$), see section 1.2.2. The vertical scale is the same to facilitate comparison between the graphs.
Figure 1.4: Ratios of expected MRUA times, $E_{\tau_N}/E_{\tau_2}$, for $N = 2, 10, 50, 100$. In all cases, as $a$ increases, the ratio approaches 1.
Chapter 1. Modeling the genealogy of a cultural trait

Figure 1.5: Change of sample size with time in model simulations (mean: dotted lines; standard deviation: gray area) and in the deterministic approximation (solid lines, equation 1.9) for three values of learning rate parameter $a \in \{1.5, 2, 5\}$ and two values of initial sample size, $X_0 = 2$ (above), $X_0 = 100$ (below), for a population of $N = 100$ individuals. Replacement rate is $r = 1$. All simulations that reach MRUA in the plotted time window are shown in light gray. One-thousand simulations were performed for each choice of parameters.

Figure 1.6: Numerical determination of the approximate relationship in (1.11). The left panel shows that $E\tau_N$ is approximately exponential in $k^*$, as defined in (1.10), $E\tau_N \approx 10^{bk^*}$. The right panel shows that $b$ in the last expression is a logarithmic function of the learning rate parameter $a$. Numerical fits have been obtained with the `lm` function of the R software, version 3.0.0, applied to log-transformed values as appropriate [83].
Chapter 1. Modeling the genealogy of a cultural trait

Figure 1.7: Times to MRUA in the multiple ancestry model (solid lines) and single ancestor model (dashed line) for an initial sample size of 2.
Chapter 2

Evolution of Cultural Systems

CITY UNIVERSITY OF NEW YORK

Abstract

Ecology, Evolutionary Biology, and Behavior

Biology Department

Doctor of Philosophy

Models and Methods in Cultural and Social Evolution

by Elliot AGUILAR
A specific goal of the field of cultural evolution is to understand how processes of transmission at
the individual level lead to population wide patterns of cultural diversity and change. Previous
models of cultural copying and innovation have assumed that traits are independent of one
another and essentially exchangeable. But culture has an architecture: traits bear relationships
to one another that affect the transmission process itself. Here we introduce an agent based
simulation model to explore the effect of cultural architecture on the process of copying and
innovation. We construct a space of all possible traits and assign them pairwise, symmetric
relationships of compatibility or incompatibility. We then implement different ways for agents
to parse these relationships, called filters. We find that introducing this simple architecture
leads to novel results. When individuals copy based on a trait’s features (here, its compatibility
relationships) they produce smaller, more homogenous cultures than when they copy based on
the cultural model. We also find that the average compatibility of a culture produced by some
filters is determined by the variance in compatibility in the space of all possible traits, a cultural
analog to Fisher’s Fundamental Theorem of Natural Selection. We discuss the implications of
considering cultural architecture in the dynamics of cultural change.
2.1 Introduction

A specific goal of the field of cultural evolution is to understand how processes of transmission at the individual level can lead to population wide patterns of cultural diversity and change. A number of authors have explored the effect of random copying as well as biased copying on various aspects of cultural change, such as the size of culture, the distribution of trait frequencies, and the rate of turnover in trait popularity [2, 12, 61]. In addition, the well-known Axelrod model [7] has investigated cultural copying in a spatially explicit context, finding cultural differentiation can result under certain parametrizations. What all these models share in common is the assumption that a cultural trait is copied independently of any other and that cultural traits are essentially interchangeable. However, even a cursory examination of human cultures shows the limitations of this assumption. For example, religions consist of sets of moral, behavioral, and metaphysical ideas that are interdependent. Similarly, a cultural object like a sword entails not only the knowledge of its manufacture and use, but also social mores about when, how, and by whom it can be used. Similar networks, often combining non-material and material elements, can be sketched for many if not all cultural phenomenon, such as views about the world, identities, social institutions, political systems and society, kinship systems, food culture, ethnicity, sex and gender, and subsistence systems. It is obvious that culture has an architecture formed by the relationships between individual cultural elements, a feature that affects both the everyday functioning of culture and cultural change. Here, we will refer to assemblages of traits and the architecture that exists across them as cultural systems. The evolution of these systems is the subject of this paper.

The idea of cultural systems is far from new. In anthropology, archeology, history, sociology and related disciplines culture and society are usually viewed as systems subject to change [18,
Chapter 2. *Evolution of Cultural Systems*  

39, 577]. One example is historical materialism (e.g. Marx 1859) which identifies production relations as central elements of society that determine political institutions, laws, customs, ideas, ways of thinking, morality, and so on. The related theory of dialectical materialism (Engels 1925) recognizes that cultural systems may contain and develop internal contradictions and weaknesses which contribute to systemic decay and promote change. A different attempt to develop a systemic theory of culture was the structuralist program that first appeared in linguistics and later in anthropology and other fields [25, 56, 68]. Its emphasis tends to be non-material culture, holding that cultural elements can only be understood within a larger, overarching system, or “structure,” typically formulated through linguistic symbols and binary oppositions (good/evil, earth/sky, and so on). Structural functionalism [26, 64, 70] viewed societies as a system of norms, customs, traditions, and institutions. Interestingly, both Lévi-Strauss and Radcliffe-Brown considered an important goal of anthropology to formally describe social structures using mathematics, but this project was never carried through. Cultural systems have also been seen as adaptations to the physical and ecological environment [14, 21, 45, 71]. In the field of history, systems thinking is for instance central to the Annales School, that considers economic, social and mental patterns, and different modes of systemic change [29, 86]. The idea of culture, and in particular non-material culture, as a system exists also in postmodern thinking (e.g., ideas about patriarchal system and structures). Systems thinking can also be found in other work on social constructions [13, 38, 75].

These verbally formulated theories about cultural systems hold powerful insights that deserve to be further developed and synthesized, and they naturally lend themselves to mathematical treatment to further sort out their assumptions, logic, predictions and explanatory power.
Importantly, a similar move towards systems thinking has occurred in genetics due to the recognition of the effect of epistasis on evolution [67]. We argue that this shift in thinking is even more important for culture; while a number of genes may be susceptible to natural selection due to the complex interactions that produce a phenotype, the relationship between cultural traits may alter the transmission process of those traits itself.

In this paper we outline a definition of a cultural system and investigate how a simple architecture impacts the evolution of that system. To do so, we present an agent based model of cultural copying where traits have underlying relationships of compatibility or incompatibility with one another. We are interested in the following basic questions about a systems view of culture:

1. How do different modes of individual transmission affect the evolution of culture (at a system level)?
2. How does the underlying architecture of relationships between traits affect the size and diversity of culture?
3. What causes cultural systems to diverge?

**Cultural architecture and properties of cultural systems**

Cultural architecture is the network of relationships that exists between cultural traits, whether the traits are already present in the population or are yet to be introduced. A simple example is provided by moral/legal system set down in the Ten Commandments. The injunctions not to commit adultery and not to covet thy neighbor’s wife are quite consistent; it is likely that belief
in one of these rules can reinforce or make it more likely to believe in the other. However, a belief in a particular rain deity would be inconsistent with the first commandment to worship only the god Yahweh. While the rules themselves (or the beliefs in them) are the cultural traits, these relationships of consistency or compatibility are the architecture across them. There are many factors that can determine trait relationships. For instance, how well traits together perform a function, how they fit into the current social environment, or how they influence the ability and motivation to invent or innovate on existing traits. Some traits are also logically consistent with one another while others are not, as in the above example of the Ten Commandments. Importantly, these relationships can be determined by factors exogenous to cultural evolution, such as physical or environment constraints; these relationships might also be determined by the outcome of previous cultural evolution itself, as in the association between graphic symbols and concepts or ideas (e.g. crosses, swastikas, etc.). Therefore, the concept of architecture being discussed should be recognized as both a guiding force and a product of cultural evolution.

We define a cultural system as an assemblage of traits and the architecture of relationships between them. Cultural systems exist both at the level of a population (i.e. all the traits present in the population and how those traits interact) and the level of an individual (i.e. one’s own traits and how they interact), roughly analogous to the ideas of a gene pool and a genome for genetic evolution.

A fundamental question is how these assemblages evolve with respect to properties such as size, diversity, and structure. To answer this question we must understand how traits affect each other’s transmission, modify existing traits, and affect the introduction of new traits. Here, we develop a representation of cultural systems in order to explore the process that molds such systems. We first assume that the diverse kinds of relationships between traits can be
summarized by the extent to which traits facilitate, hinder, or are indifferent to each other’s acquisition. We will also assume that each trait’s effect on another trait is independent of other traits. This is a simplifying assumption and such independence is not always the case. For example, knowledge of how to fish and knowledge of how to grow crops may lead to acquiring the practice of using fish carcasses as fertilizer. However, it is unlikely that this practice would arise from either of the other two in isolation.

Let $R(i,j)$ be a binary relation on the set of traits that describes the relationship of trait $i$ to trait $j$. This binary relation may take the values of 1 (compatible), -1 (incompatible), or 0 (indifferent), thus $R : (i,j) \rightarrow \{-1, 0, 1\}$. The idea here is that the complex interactions between any two traits can be reduced to one of three possibilities, that traits aid (compatible), hinder (incompatible), or are indifferent to each other’s transmission. We further assume that the relation is commutative, so that $R(i, j) = R(j, i)$, but not necessarily transitive (i.e. $R(i, j) = 1$ and $R(j, k) = 1$, does not imply $R(i, k) = 1$). Of course relationships may exist between any number of traits, and their effects need not assume discrete values. Similarly, relationships may be fundamentally asymmetric, for example language must be acquired prior to literacy, but not vice versa. While we recognize these complexities, we start with these simplifying assumptions in order to facilitate our investigation.

Whatever relationships may exist between traits, they only matter insofar as they affect copying and innovation. There are many ways that these relationships could be parsed or incorporated in cultural transmission, and we refer to these as cultural filters. These filters can intervene in cultural transmission in many ways. For example, they may affect how individual’s make decisions about what and when to copy; or they may affect what individuals make available for copying, through teaching or demonstration; finally, they may affect which traits are invented
Chapter 2. Evolution of Cultural Systems

and by whom. We implement examples of each of these kinds of filters in the model described below.

2.2 Model description

We constructed an agent-based simulation in the Python 3 language to explore the evolution of cultural systems. In our simulations, each agent possessed a repertoire of traits and could acquire more traits by copying from a partner in repeated pairwise interactions, or by introducing a new trait. All possible traits were designated at the start of the simulation in what we termed the trait universe.

2.2.1 Trait Universe

The trait universe defines the set of all possible traits that can exist in the population and the values of $R(.,.)$ for all pairs of traits. This construction provides a simple architecture that can be varied along a single dimension: the proportion of compatible trait pairs. Varying the composition of the trait universe captures the notion that for different domains of culture, the range of possibilities for cultural evolution will vary. Consider the morphology of a language: the diversity of human languages demonstrates the enormous range of possibilities that exist for combining morphological, syntactical, and phonological elements. By contrast, a technological system, such as a hand axe, will place more constraints on which elements can be successfully combined, for example different methods of binding the axe head to shafts of different materials. The trait universe allows us to represent these differences in design space by varying the number of compatible trait pairs. Let $0 \leq c \leq 1$ be the measure of how restrictive (low $c$) or permissive
(high $c$) the trait universe is. To generate a universe of $T$ traits we randomly assigned the values of $R(.,.)$ for all $T^2$ pairs, excluding all such pairs $(i,i)$ for which $R(i,i) = 1$. The remaining pairs are compatible with probability $c$ and incompatible with probability $(1 - c)$. In our analysis we examine the effect of varying $c$ on the outcomes for cultural systems.

### 2.2.2 Filters

To understand the effects of cultural architecture, the compatibility relationships delineated in the trait universe must have some affect on the introduction and transmission of traits in the population. In our simulation we assigned agents decision rules that take into account the relationships between traits to calculate the probability of an action, such as copying or innovating. We call these decision rules, filters. We constructed a filter for each of the following basic actions:

1. **What and when to copy**: When agents met in pairwise interactions, they selected one of their partner’s traits for potential copying. We allowed agents to copy one another based on either the compatibility of the trait they may acquire with their existing repertoire (Trait Filter); or the overall compatibility of the repertoire of their partner in the interaction with their own (Model Filter).

2. **What and when to innovate**: We allowed an agent to decide whether to introduce a new trait from the trait universe based on its overall compatibility with the agent’s existing repertoire (innovation filtering).

3. **What to display**: We allowed agents to decide what traits to make available for copying by others according to the compatibility of traits within their repertoires (Display Filter).
Chapter 2. *Evolution of Cultural Systems* 31

The trait, model, and Innovation filters all calculate the probability of acting \( (p) \) according to the logistic function:

\[
p = \frac{1}{1 + e^{-ks}}
\]  

(2.1)

where the variable \( s \) is the score, a value that incorporates the compatibility between traits, and the parameter \( k \) determines the strength of the dependence on \( s \) (varying \( k \) did not alter the results qualitatively, and in the following results \( k = 1 \) for all runs). The score was calculated differently for the trait, model, and IFs, in the following ways:

1. **Trait filter (TF):** \( s \) was the average compatibility of the new trait with the existing traits in the agent’s repertoire.

2. **Model filter (MF):** \( s \) was the average compatibility between both agents’ trait repertoires.

3. **Innovation (IF):** \( s \) was the average compatibility of the new trait with the existing traits in the agent’s repertoire.

Finally, \(-1 \leq s \leq 1\), as the underlying compatibility relationships assume values of 1 or \(-1\).

The Display Filter (DF) worked differently. Here, an agent calculated the average compatibility of each trait in her repertoire with all others in the repertoire, and assigned a probability mass function to all her traits according to these compatibilities. She then selected a trait to be made available for copying in an interaction according to these probabilities.

To isolate the effects of the different filters, each filter was implemented independently. When the trait or MF was in effect, agents displayed traits at random (uniformly) and were allowed to innovate with a fixed probability that led to, on average, one innovation per lifetime. When
the IF or DF were in effect, copying between individuals occurred with a fixed probability of 1/2.

2.2.3 Cultural Copying Model

In our simulation model, agents in a population of fixed size, \( N \), encountered one another at random at discrete time intervals. In each encounter, an agent made the decision to copy a cultural trait according to a probabilistic decision rule determined by the filter(s) in effect in the particular simulation; they were then potentially able to innovate and/or be selected for replacement (death) by a new individual with no traits.

To summarize, the simulation model proceeds in the following steps:

*Initialization*: a trait universe of \( T \) traits is constructed. All pairs of traits are assigned compatibility relationships \( \{1,-1\} \) at random according to a specified proportion \( c \) of compatible relationships.

*Simulation*:

At each time step:

1. Agents select another agent at random from whom they may potentially copy a trait.

2. Agents choose to copy one of their model’s traits according to a decision rule specified by the filter (in the absence of trait or MF, probability of copying is 1/2). If the DF is in effect, then this will determine which traits are available for copying in an interaction.

3. Agents are given the opportunity to innovate (directly sample a trait from the trait universe) such that each agent will introduce a trait on average once per lifetime. However, if the IF is
in effect, the probability of introducing this trait will be determined by its compatibility with their current repertoire (see Appendix D).

4. Each agent is selected to die and be replaced with a naive individual with probability $\frac{1}{N}$.

### 2.3 Results

In our simulations agents introduced traits (from the trait universe) and transmitted them to one another via copying. This resulted in the emergence of systems of cultural traits in the simulated populations that varied in size, diversity, average compatibility, and other measures. We wanted to know how these systems would evolve with respect to the two main variables: the average compatibility in the trait universe and the particular filter in effect for parsing compatibility relationships. In the following results we varied the average compatibility in the trait universe ($c$) and ran each filter in isolation for populations of fixed size ($N = 100$) for $1 \times 10^4$ rounds of interaction. The average lifespan of an agent was 100 interactions and the trait universe contained $T = 1 \times 10^4$ traits. We ran ten simulation runs for each constellation of parameters.

For comparison we also simulated runs in which no filters at all were in effect, and copying, innovation, and display all occurred at random. Importantly, agents in these ‘neutral’ runs did not at all take compatibility relationships into account. Thus, the systems that emerged in these runs provide useful benchmarks for the systems that evolved under the various filter types.

Time series of all the cultural system measures showed that they reached stationary values well before the end of the simulations. Thus, for each measure we report the time averages of the
final 20% of simulation runs. We directly address our results to the questions posed in the introduction.

2.3.0.1 How does the underlying architecture of relationships between traits affect the size and diversity of culture?

We measured the size of culture as the number of cultural variants present in at least one member of the population. We also measured the average number of traits possessed by a single individual (repertoire size). Fig. 2.2 shows the culture and repertoire sizes for all run types compared to the neutral case. In trait universes with low compatibility \((c < .5)\) the filter runs produced smaller cultures than the neutral case, while in higher compatibility universes filter runs produced larger cultures (with the exception of the display filter). This pattern follows our expectations, since low \(c\) values will lead to lower probabilities of copying/innovation than in the neutral case—for high values, the converse is true. The pattern is the same for repertoire sizes, where filter runs produced smaller repertoires in low \(c\) universes and larger ones for high \(c\) when compared with the neutral case. Thus, more permissive (high \(c\)) architectures lead to larger cultures than would be expected in a neutral "bean bag" model, while less permissive architectures had the opposite effect.

The Trait and MFs both affect agents’ decisions to copy. It make sense to compare their effects on cultural systems directly. In Figure ??a we see that TF leads to smaller cultures than MF in the low \(c\) universes, while for the high \(c\) universes there is essentially no difference in size. At the same time, the Trait filter repertoire sizes are larger in the low \(c\) universes (see Figure ??b). We also find that individuals are far more similar and more compatible in the TF runs for low \(c\) universes (see Figures ??c and ??d). The results can be summarized thus: TF leads to
smaller cultures, where individuals have larger repertoires and are more similar to one another than under MF.

The above result may seem counterintuitive; shouldn’t bias in favor of the model overall make individuals more similar to their cultural models, and thus, increase similarity, as opposed to cherry-picking for compatible traits? The reason for this finding can be understood by considering the fate of a novel innovation in both cases when \( c \) is low. The rate of innovation is the same in both cases (i.e. \( 1/(\text{generation length}) \)). However, under TF, an innovation is less likely to spread, since it is likely that it is not compatible with a copier’s traits due to the overall low compatibility in the trait universe. Innovations that do not propagate in the population are more likely to be lost due to death events because of their low frequencies. Therefore, the overall culture size remains small, since new variants are consistently being lost. The rare traits that possess many compatibility relationships will be highly favored for copying and will spread through the population. As a result, TF causes individuals to accumulate these particularly compatible traits, leading to larger repertoires that are highly similar between individuals.

Under MF an innovation has a better chance of spreading since its compatibility score averages over all the other traits a model possesses. This means that innovations are less susceptible to loss through population turnover and culture size can grow larger. However, highly compatible traits don’t have as great an advantage as they do under TF, since again, their compatibility score is now being computed by taking into account all the model’s other traits. Repertoire sizes remain smaller since copying is less frequent due to the low compatibility in the trait universe and individuals become less similar. The differing effects of these two filtering mechanisms reflects the importance of cultural architecture for the outcomes of cultural copying and innovation.
For high $c$ universes, the innovation filtering led to the largest cultures of any filter, as the rate of introduction of novel traits peaked. By contrast, there was little effect on repertoire size and sharing compared with the neutral case. This makes sense, given that this filter only affected which variants were introduced in the population, and not copying, which occurred at random.

Among the filters, DF led to the largest culture sizes in low $c$ universes. While copying is still neutral, DF favors variants that are highly compatible just like TF. However, as some variant must be displayed, and copying is unaffected, repertoire sizes remain close to the neutral runs. Sharing is lower than under either trait or MF and similar to the IF, which acts infrequently.

In the high $c$ universes, the DF differs little from the neutral case for all measures. As the variance in the number of compatibility relationships traits possess is reduced, the probability of any one trait being displayed approaches a uniform distribution across the model’s repertoire of traits; the system approaches random display/copying.

We have seen that cultural architecture has a marked effect on the size of culture, and the similarity between individuals.

2.3.0.2 How do different modes of individual transmission affect the evolution of culture at a system level?

We measured overall differences in the cultural systems that arose under each filter with respect to three measures of compatibility: 1) the average compatibility among traits within an individual’s repertoire, 2) the average compatibility between individuals’ repertoires, and 3) the average compatibility of all traits present in the culture at the population level.
Fig. ?? shows these three compatibility measures for each filter with the neutral copying as a baseline (i.e compatibility of filter run minus compatibility of neutral run). All differences are positive, since all filters had the expected effect of increasing compatibility over the neutral case. This makes sense since the filters incorporate compatibility into the copying/innovation process, and therefore ‘select’ for compatibility.

Over most of the range of $c$, TF produced the greatest gain in compatibility with respect to the neutral run, followed by the DF, MF, and IF. As we have already seen, the Trait and DFs have similar effects, though they differ in magnitude. The compatibility results are consistent with the results above, in that the filters that lead to increased similarity also show the highest compatibility values.

For all filters, the difference in compatibility compared to the neutral runs must go to zero at $c = 1$, since all traits then are compatible. What is striking is how quickly the compatibility difference declines for the TF, and, to a lesser extent, the DF. To get a better intuitive understanding of why this occurs, we must look at the effect of varying $c$ on the trait universe itself. The process for generating the trait universe is equivalent to an Erdos-Renyi random graph process, where nodes are traits and edges indicate positive compatibility relationships (i.e. $R(i, j) = 1$). The degree of each node is just the number of positive compatibility relationships a particular trait possesses. From the standard results for the degree distribution on such graphs, the coefficient of variation (i.e. the ratio of the variance to the mean) of the degree is linear in $c$. In Fig. 2.3 we plot the compatibility difference for each filter against the coefficient of variation in degree for trait universes with difference $c$ values. We see that the trait and DFs, the two that showed the steep declines in compatibility differences with increasing $c$, responded most—that is to say, produced the greatest compatibility gains over the neutral runs—when the
coefficient of variation in degree was highest. If we recall that the degree of a node is just the number of compatible relationships a trait possesses, and that the more compatible relationships a trait has the the higher the number of individuals that are likely to copy it, we can see that degree is a proxy for the fitness of a trait. The coefficient of variation in degree, $\frac{\text{var}(d)}{\bar{d}}$, then looks very similar to the term $\frac{\text{var}(w)}{\bar{w}}$ in Fisher’s fundamental theorem of natural selection. Fisher’s theorem states that the change in the mean fitness of a population ($\bar{w}$) due to selection will be equal to $\frac{\text{var}(w)}{\bar{w}}$. Thus, we see that the TF and DF act like selection forces, increasing the compatibility of the culture (and thus, the average repertoire size of individuals) as $\frac{\text{var}(d)}{\bar{d}}$ increases.

Why then do MF and IF not behave similarly? As discussed above in the results on size, the MF mitigates the advantage of traits with a high number of compatible relationships as well as the disadvantage of traits with a low number of compatible relationships. This makes the MF a more diffuse selective force with respect to creating more compatible systems than the neutral copying, which leads to a lower compatibility for each measure, and a slightly different response to $\frac{\text{var}(d)}{\bar{d}}$. The IF only affects the systems by introducing variants that may be more compatible, but once a variant enters the population its probability of being copied is fixed, no matter how many compatible relationships it has. Therefore, IF also has a weaker ‘selective’ effect with respect to compatibility.

We have seen that when an architecture is present, TF, which targets specific traits, exerts a greater selective pressure than MF, and produces more compatible cultural systems. While DF and IF also act on traits directly, DF has a greater effect in selecting for compatibility because it affects the copying process directly, while IF only affects what variation in traits is available for being selected.
2.3.1 What causes cultural systems to diverge?

We were interested in whether the compatibility relationships between traits in the trait universes would steer subsets of the population toward developing consistently different cultural systems from one another. We imagined this process as a cultural analog of speciation, whereby certain groups of individuals would be largely similar to one another while significantly different from others. To search for evidence of this kind of cultural clustering we examined the distribution of nearest neighbors in each cultural system. For each agent, we calculated the percentage of shared traits with all other agents, then ranked these “neighbors” from highest to lowest according to these values. Averaging over all agents for each rank produced the nearest neighbor distribution (NND). The NND can be qualitatively examined for evidence of clustering. If the change in similarity between individuals is clinal, then the NND will be smooth; where there are discrete boundaries of difference between individuals, we should expect the distribution to be somewhat flat initially, indicating high similarity among nearer neighbors, and then decline steeply at the domain boundary. While imperfect, the NND allowed us to look for salient evidence of clustering. We found no strong evidence of clustering in any of our runs. For all runs the NND was either concave or linear, with the notable exception of TF for $c = .1, .3$ (see Figure 2.4). However, we caution against overinterpretation of this difference; it is likely due to the high degree of similarity observed in the TF runs. Thus, with the architecture we implemented we do not observe cultural divergence or speciation.
2.4 Discussion

We propose taking a systems view of culture as a next step in the development of theory in cultural evolution. We believe that in order to understand the differences between cultural assemblies, as well as how different assemblies emerge from previous ones, the relationship between traits must be considered. In order to explore the consequences of this view, we constructed an agent-based simulation of cultural copying. Our simulations implemented the idea of an architecture across cultural traits (in the form of pairwise relationships of compatibility and incompatibility) and various ways for that architecture to influence cultural copying (filters). We then examined how different ways of parsing these relationships affected the process of change in the overall cultural system.

2.4.1 Filters

The filtering mechanisms all produced more compatible cultures than those produced by random copying. However, the size of this effect, for some filters, was determined by the variance in compatibility in the trait universe, a proxy for the variance in fitness among traits. For TF and DF we showed a relationship that seemed roughly analogous to Fisher’s Fundamental Theorem of Natural Selection,

\[ \Delta \bar{w} = \frac{\text{var}(w)}{\bar{w}} \]

where \( w \) is fitness and \( \bar{w} \) is population mean fitness. Fisher’s theorem states that the change in the population mean value of fitness due to natural selection is equal to the population variance in fitness [31]. In our cultural example, it is traits that are replicating (i.e. ‘reproducing’); the more compatible relationships a trait possesses the more likely it is to get copied. Therefore,
the number of compatible relationships a trait possesses can be taken as a proxy for the trait’s ‘fitness’, the number of individuals in whose repertoires the trait ends up. However, the analogy with Fisher’s Theorem is imperfect, as the relationship between average compatibility and variance in compatibility in the trait universe is not linear (see Fig. 2.3). This is because a traits’ average compatibility with other traits is not the only factor that determines its fitness. The filter being used by the agents along with the trait compatibilities together determine the fitness landscape for traits.

Taking the trait’s point of view, our model can be viewed as a social evolution analog of trait evolution. Each cultural trait’s fitness is determined by the relationships it has with other traits in the trait universe. The traits in an agent’s repertoire can be seen as a ‘group’, whose interactions with other groups determines the fitness of the members of that group, the traits themselves. This is an insight which might prove profitable in developing analytical models of cultural systems.

TF and MF showed opposite effects on culture size, repertoire size, similarity and cultural compatibility. Trait filtering led to smaller, more homogenous cultures with larger repertoires than MF. Aside from the counterintuitive result that ‘cherry picking’ traits increased similarity between individuals when cultural architecture was taken into account, there is another interesting consequence of these different behaviors. Cultural traits possessed by only a few individuals may have enormous consequences for the population at large, for example, only a few members of a band of hunters may be skilled in making arrowheads, but all hunters rely on this knowledge. It is therefore reasonable for some domains of culture to use the culture size—the number of distinct variants—as the relevant measure of cultural diversity. By this measure, MF produced
the more diverse cultures in our simulations than TF. Given the small repertoires and low similarity under MF, the loss of some fraction of the population is likely to cause a greater loss to cultural diversity than in TF. This suggests that for cultural domains that are dominated by either form of copying, we should expect different magnitudes of cultural loss in the event of a population shrinkage. Potentially, this can be explored empirically in historical examples of cultural loss, such as occurred after the fall of the Western Roman Empire.

2.4.2 Model assumptions and future directions

Our model assumed that the architecture across all cultural traits was delineated and fixed at the start of the process. In reality, new traits and new compatibility relationships arise over time and as a result of historical contingency. Thus, in reality this architecture will be both a force and a product of cultural evolution. For example, a swastika and and Star of David are unlikely to be considered compatible symbols, not because of some a priori nature of their meanings or appearances, but because of a particular history of cultural associations attached to both symbols. While we recognize this fact, constraining the architecture at the outset of the simulations allowed us to examine the effect of cultural architecture more directly; by varying $c$ we could explore a range of scenarios for cultural copying. Additionally, our assumption of pairwise, symmetric relationships between traits was simplistic. In the future, asymmetric relationships, which might indicate sequential learning of traits could be explored and allowing relationships between triads, or higher number tuples of traits would lead to more complex architectures.

We did not find evidence of clustering or cultural speciation in our model. We assumed the population of agents was freely mixing. Imposing a network structure on agent interactions
would certainly have made observing cultural speciation more likely. However, these networks of cultural copying may themselves be the product of cultural evolution (e.g. previous cultural speciation), and we believe it is an important goal to see how the cultural evolutionary process can construct such networks. Our trait universe architecture was equivalent to a network structure for cultural traits themselves. While this simple architecture did not result in speciation, we believe more sophisticated architectures (perhaps those including historical contingency) may produce clustering and we hope to explore this idea in future work.

Finally, all of our cultural traits were neutral, in the sense that they had no effect on survival and performed no function whose performance could be assessed by agents. This last point is likely to be a very strong factor in governing the architecture of culture.

Our model was an attempt to formalize the concept of cultural systems. Researchers in cultural evolution have revealed a number of important phenomena using simple copying models. These approaches, often inspired by population genetics, are quite different from the traditional views of cultural anthropologists and other students of cultural change. We hope that incorporating a system-wide view will help bridge the gap between cultural evolutionists and cultural anthropologists, and hopefully lead to new insights into cultural change.
Chapter 2. Evolution of Cultural Systems

Figure 2.1: (A) Average culture size (across ten runs) with increasing $c$ in the trait universe. (B) Average repertoire size (across ten runs) with increasing $c$ in trait universe.
Figure 2.2: Average compatibility of cultural systems (A) within individuals, (B) between individuals, (C) population-wide.
Figure 2.3: Filter-Neutral compatibility vs. coefficient of variation in number of compatible trait relationships.

Figure 2.4: Nearest neighbor distribution under TF and MF in universe with c=.1
Chapter 3

An evolutionary model of contact induced language change

Abstract

Language shift is when a group of speakers adopts a new language, and occurs as part of the larger phenomenon of language contact. It has long been observed that language change is accelerated during shift situations. The standard explanation for this accelerated change has been the introduction of novel linguistic forms by new speakers during the second language acquisition (SLA) process. This hypothesis is based on historical reconstructions of contact situations and has never been formally tested on empirical data. In this paper, we construct an agent-based model to formalize the hypothesis that L2 speakers are responsible for accelerated language change during language shift. In our simulations, a population experiences demographic change via the birth of native (L1) speakers, recruitment of L2 speakers, and death. However, only L2
speakers have the potential to ‘mutate’ (introduce a new variant) on entering the population. We then parameterize the model using demographic data from Maputo, Mozambique—where rapid shift from Bantu-languages to Portuguese has been occurring for the past forty years—and compare our model predictions to a rare diachronic data set on two linguistic features of Portuguese in Moputu. We find that our basic model is a poor fit to either data set. Next, we modify the model by allowing L2 speakers to introduce a novel variant at any point during their first five years in the population, a feature that represents the fact that the SLA process is not instantaneous (we find support for the five year duration in the literature). We find that the extended SLA model is a good fit to one of our datasets—we discuss plausible reasons for why the other data set is such a mismatch. Finally, we discuss typological differences between contact-induced and non-contact-induced language change and suggest that multiple introductions of a new linguistic variant by different individuals may be the mechanism by which SLA accelerates language change.

3.1 Introduction

Language shift occurs when a group of people adopts a new language. History is rife with examples of language shift, such as the spread of vulgar Latin across the Roman Empire in Western Europe, or the adoption of Arabic during the Muslim conquests in the Middle East and North Africa. While all languages change over time, language shift is widely believed to accelerate the pace of language change. The standard explanation for this increased change is that new speakers introduce novel linguistic forms at a greater rate than do native speakers as a result of the second language acquisition (SLA) process. These novel forms are then transmitted to later generations of both native (L1) and second language (L2) speakers causing the language
to change [84, 88, 89]. If true, this idea predicts that the pace of linguistic change should be correlated with the rate of introduction of L2 speakers.

The previous hypothesis is largely based on historical reconstructions of shift situations where data on the number of native (L1) and second language (L2) speakers is scant or nonexistent. Recent studies based on synchronic linguistic data have shown negative correlations between proportions of L2 speakers and morphosyntactic complexity as well as lexical diversity (Bentz et al., 2013; 2015). While these results support the hypothesis that SLA is involved in contact-related change, the lack of diachronic quantitative data and the difficulty of setting up experiments means that there is little information about the mechanism by which SLA accelerates change. In this sense, the hypothesis has never been formally tested.

Simulation models are a useful way to explore the mechanisms by which new linguistic forms spread in a population through innovation and interaction, overcoming the difficulty of scarce data and empirical constraints. Mathematical and computational models have been employed in the field of language evolution to explain a number of phenomena, for example the emergence of grammatical or phonological systems in human language [24, 53, 77, 78]. Simulation methods can be combined with empirical data to test the explanatory power of models. For example, Baxter et al. introduced a model for contact-related change whose predictions were compared to demographic and linguistic data from the evolution of New Zealand English [10]. more recently, Jansson et al. presented a model for the formation of creoles, another contact-related phenomenon, and applied their results to the case of Mauritian Creole [47]. Using agent-based simulations and demographic data from the colonization of Mauritius, Jansson et al. were able to accurately model how speakers of multiple languages converged on the current structure of Mauritian creole.
In this paper, we use simulations to model language change during language shift. We modify the creole formation model of Jansson et al. to allow for both the introduction and propagation of novel variants in a population of L1 and L2 speakers of the same language. In order to test the hypothesis that variation is introduced during the second language acquisition process, we allow only L2 speakers to introduce the novel variant while entering the population. We then simulate this model in a fixed and an expanding population. Next, we run the model with demographic parameters based on the growth of the Portuguese speaking population of Maputo, Mozambique in the years 1975-2007, during a sustained period of contact between Portuguese and Bantu languages and compare the model results with data on changes in two grammatical features of Mozambican Portuguese. We test two models, one simpler model where individuals may introduce the novel variant only at the very moment of entering the population, and another where the second language acquisition process lasts five years allowing individuals to introduce novel variants on multiple occasions. Our results show that the second model is a good match to the data on change in preposition forms in Maputo Portuguese, while it is a poor match for the data on verb forms. In the discussion we address some plausible explanations for this mismatch.

3.2 Methods

3.2.1 Model of language change

Our aim was to model the evolution of a linguistic feature due to repeated interactions in a group of native (L1) and second language (L2) speakers, where only the latter group could introduce novel variants into the language. For example, if the linguistic feature is subject-verb
agreement, as in “I/you/we/they go; he/she/it goes” we can imagine that some L2 speakers, while acquiring the language, reduce this paradigm to the simpler “I/you/we/they/he/she/it go” (Note: this is actually similar to what happens in Mozambican Portuguese, and one of our data sets is on reduced subject-verb agreement, see below). In this case, the loss of the verbal suffix represents the novel variant introduced by L2 speakers, while the maintenance of the same element represents the conservative variant. Thus, introducing a novel variant in the model does not necessarily mean adding an element to the language in question, but rather changing something, which may imply adding, removing or replacing a linguistic element. The frequencies of usage of the conservative and novel variants then change over time due to repeated interactions between speakers.

To model the interactions between speakers, we adapted the model of linguistic interaction introduced in Jansson et al. [47] to model creole formation. This model allows individuals to interact and update their knowledge according to the outcomes of interactions. For a given language feature, an individual, i, is characterized by a discrete usage distribution, \( p_i := \{p_{i1}, p_{i2}, \ldots, p_{in}\} \), where \( p_{ik} \) is the probability of using variant \( v_k \) in an interaction, and \( \sum_{k=1}^{n} p_{ik} = 1 \). At discrete time steps, a round of interaction occurs: each individual encounters another and both choose to utter a variant according their respective usage distributions. In the encounter between agents \( i \) and \( j \), let the variant uttered by \( i \) be \( u_i \), and the variant by \( j \) be \( u_j \). After an interaction, an individual \( i \) updates her probability of using each variant in the following way,

\[
\begin{align*}
    p'_{ik} &= p_{ik} + (1 - p_{ik})l, \text{ if } u_j = k \\
    p'_{ik} &= p_{ik} - p_{ik}l, \text{ if } u_j \neq k
\end{align*}
\]
Here, $0 \leq l \leq 1$ is the learning parameter that determines the importance an agent gives to each interaction; thus, high $l$ means that agents change their probabilities of usage substantially after just one interaction, while low $l$ means that they update only slightly. In the absence of innovation the model results in convergence on a stable form of the language.

### 3.2.2 Demographic change/growth

We allowed demographic change to occur through births of L1 speakers, recruitment of L2 speakers, and deaths of both types. Each newborn L1 selected two agents at random as ‘parents’ and averaged their usage distributions to obtain its own. This represents the idea that native speakers will enter the population with a limited number of linguistic models (e.g. their two parents). Each recruited L2 speaker averaged the entire population to obtain its initial usage distribution, representing the idea that a second language learner will interact with many individuals when first acquiring the language. However, with probability $\mu$, instead of averaging the population, a newly recruited L2 speaker assigned the novel variant a probability of 1 and all other variants a probability of zero. This allowed L2 speakers to introduce novel variants on entering the population. We set $k = 1$ to be the existing variant, and initialized the simulation with $p_{i1} = 1$ for all agents to represent the idea that initially only the standard variant is present.

We allowed agents to complete $T$ rounds of interactions before demographic change occurred. In the fixed population model, $dN$ individuals are selected for death and replaced with either a “birth” (producing a L1 speaker), “recruit” (producing a L2 speaker), according to the rates $b$ and $r$, respectively. In the case of the expanding population simulations, deaths, births, and
recruitments all occurred with rates $d, b$, and $r$, respectively. The overall growth rate was given by $b + r - d = .05$ for all expanding population runs.

We briefly note that the model described above is neutral in the sense that interactions do not favor either variant. To demonstrate this, let $\bar{p}_k = \frac{1}{N} \sum_{i=1}^{N} p_{ik}$, the population mean use of variant $k$. The expected change in the average frequency of usage of the variant is then $E[\bar{p}_k' - \bar{p}_k] = 0$, (see Appendix E). However, extinction is guaranteed for the existing variant, since $\bar{p}_1 = 0$ is an absorbing state, while the novel variant can always be rescued from extinction by recurrent mutation. In the rest of the paper we will be concerned with the rate at which the novel variant increases in mean frequency of usage.

We ran simulations for 100 years, with 365 interactions per year, with demographic change occurring at the end of each year. For the comparison with language change data from Maputo, Mozambique, we ran simulations based on demographic data from the years 1975-2007 (see Appendix F).

### 3.2.2.1 Maputo language change

The Republic of Mozambique achieved its independence from Portugal in 1975. During the colonial period, use of the Portuguese language was restricted to Portuguese colonizers and Afro-Portuguese urban dwellers, and was spoken by few other Mozambicans [30]. After independence Portuguese was adopted as the only official language. Due to the expansion of the educational system and the migration of Bantu-speaking rural Mozambicans to cities in large numbers, usage of Portuguese exploded during the years following independence. The number of speakers of Portuguese is still increasing throughout the country, and especially in the capital Maputo.

We presume the novel variants to be nonexistent at the start of the simulation in 1975, as the
spread of Portuguese among Mozambicans in Maputo has mainly occurred since Mozambique’s independence in 1975. Thus, linguistic features in Maputo Portuguese that are not registered in European Portuguese may be considered to have appeared after Mozambique’s independence, or shortly before it. The Maputo setting provides an opportunity to study an ongoing language shift where new linguistic variants have spread during a relatively short period (see Appendix F). Two comparable data points from 1997 and 2007 combined with the assumption that the conservative variant was used exclusively before Mozambique’s independence provide us with rare diachronic data on linguistic variation in a contact/shift situation. We chose to focus on two grammatical features that are registered in both datasets: reduced subject-verb agreement and innovative preposition use including the addition, exclusion or replacement of a set of Portuguese prepositions. We chose one morphological feature (verbal agreement) and one syntactic feature (preposition use) rather than lexical or phonological ones, considering that morphosyntactic features are generally more resistant to change than others [74]. This implies that the registered change in these features during a restricted period of time would probably not have occurred without the influence of language contact. For both features, the use of the novel variant increases during the period. For the reduced verbal agreement, the novel variant, consisting of the reduction of the third person plural suffix, starts at a presumed zero in 1975, spreads to a level of 13.2 percent in 1997 and reaches 26 percent in 2007. Innovative use of preposition is also presumed to be inexistent in 1975, affecting 20 percent of the registered variable contexts in 1997, increasing to 28 percent in 2007. Furthermore, we have data on the number of L1 and L2 speakers in Maputo from censuses in the years 1980, 1997 and 2007; we can estimate these numbers for the year of independence, 1975, based on population data from the World Bank, and L1/L2 speaker estimates for 1952, 1955 and 1970 made by Firmino. Thanks to this unique constellation of data, Maputo is one of the best empirical examples for testing a model
of language shift and change.

Table 3.1: Demographic data on Portuguese speakers in Maputo, Mozambique.

<table>
<thead>
<tr>
<th>year</th>
<th>L1 speakers</th>
<th>L2 speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975</td>
<td>100*</td>
<td>20,000</td>
</tr>
<tr>
<td>1980</td>
<td>6,525</td>
<td>326,521</td>
</tr>
<tr>
<td>1997</td>
<td>241,709</td>
<td>599,438</td>
</tr>
<tr>
<td>2007</td>
<td>470,690</td>
<td>612,992</td>
</tr>
</tbody>
</table>

*Due to the massive emigration of Luso-Mozambicans around the time of independence, the number of L1 speakers in 1975 was estimated based on overall population growth and the recruitment rate.

3.3 Results

We were first interested in how the model parameters $r$ (recruitment rate of L2 speakers), $\mu$ (innovation rate), and $l$ (learning parameter) affected the spread of the novel variant in the population. For both the fixed and growing population models we ran the ten simulations at each constellation of parameter values for 100 model years (36500 rounds of interaction). We found that the rate of increase of the novel variant was most sensitive to $\mu$ and $r$, whose product determined the rate of introduction of the novel variant. The learning parameter $l$, did not significantly affect the outcome of simulations beyond leading to increased variability across sets of identically parameterized runs. In the following results we use the lowest learning rate of .05 for the following two reasons: 1) we expect the actual weight given to any one encounter to be small and 2) the lower learning rate means lower variance in the runs and sets a higher threshold for fitting the model to data.

Fig 3.1 shows the means of ten runs for each constellation of parameters. A notable difference between the fixed and expanding population runs is that the latter showed more stable mean trajectories. This results from the fact that for every model year after year 1 the expanding
population models have a larger population, effectively increasing the sampling size of each run and reducing the variance across runs. This sampling effect is similar to that of genetic drift; for small populations the change in the mean use of the novel variant is likely to be more erratic.

Next we compared our runs parameterized by the Maputo demographic data to our usage data on Maputo Portuguese. We computed the 95% confidence intervals for each ensemble of runs (assuming the value of the mean frequency of usage for each run was drawn from a normal distribution, which was a good qualitative fit). We considered a given parameter setting a fit to one or both of the data sets if it included the 1997 and 2007 data points in its 95% confidence intervals for those same years.

Fig. 3.2 shows that none of the model runs met the criterion above. Generally, the model predicted a reduction in the rate of spread of the novel variant in each time period, as the rate of recruitment declined as well. Taking this into account, we made a modification to the initial model. Instead of only allowing incoming $L_2$ speakers to introduce novel variants, we allowed all $L_2$ speakers to spontaneously mutate at the start of each year for the first five years after they first entered the population. This modification more closely resembles the reality of the SLA process, which occurs over a number of years. Furthermore, empirical findings on the effect of length of residence in SLA indicate that five years is a good estimate for the duration of SLA [4, 22]. Except for this modification, the model remained unchanged. The new results show that the model runs do fit by the above criteria) for the preposition use data (see Fig. 3.3). However, the verb forms, which seem to show an increased rate of change in the period 1997-2007, still did not fit the model predictions, as there were no parameter settings that met the fit criterion.
3.4 Discussion

This paper represents the first time that a hypothesis about the mechanism that results in accelerated language change in language shift contexts has been formalized and tested. We modeled the introduction of novel variants due to an influx of L2 speakers in both a fixed and an expanding population. We then compared the model behavior with empirical data collected in Maputo, Mozambique in 1997 and 2007. The simpler model, where L2 speakers could only introduce a new variant on entering the population, showed a strong decline in the rate of spread of the novel variant after 1980, due to lower population growth after this year. This model could not account for the changes observed in the data on verb forms and preposition usage, though the preposition data did show a slight decline in rate of spread in qualitative agreement with the model runs. When we extended the period of second language acquisition to five years in the model, and let the L2 speakers introduce the new variant on multiple occasions during this period, the model did account for the spread of prepositional innovation. The data on the reduced verbal agreement showed an increase in the usage of novel variants that remained outside of the model predictions. The low number of data points and the relatively small size of our datasets makes it difficult to draw any strong conclusions concerning the validity of the model predictions. However, while the simpler model seems to be insufficient to account for the contact induced-change observed in Maputo, we do observe a relatively good match between the second model, including the extended period of SLA, and our preposition data. This is a promising result given that we only constrained the model by the demographic change of Maputo Portuguese speakers.

It must be reiterated that our model assumed neutral evolution of the linguistic form. Given this assumption, the departure of the verb data from model predictions may be indicative of directed
or selectional forces acting on this feature. For instance, the new verb forms may be easier to learn and/or produce (the conservative form is longer and the distinction between singular and plural form for third person implies a more complicated verbal conjugational paradigm) and so spread due to a bias among speakers to be more economical with their articulatory/cognitive effort. Learnability of a language has been shown to increase particularly in growing speech populations with many L2 speakers [43]. On the other hand, the preposition forms might be no more or less difficult to learn or produce than the existing variants and thus evolve more neutrally. However, these points are speculative without more study of the specific linguistic forms, which falls outside the scope of the present study.

One limitation of our model is that it assumes that the learning parameters are the same for all individuals. This may not be true in the sense that L1 and L2 speakers may perceive one another’s different speaker status and respond differently in mixed interactions than in interactions among their own class of speakers. Different speakers may also have different biases towards conservative and novel variants, depending on factors such as their age or their L1/L2 status. Future work should explore the importance of varying learning parameters for interactions between the different speaker classes, as well as heterogeneous biases in a contact situation. Additionally, there may be heterogeneity in learning parameters across all individuals, regardless of speaker class.

Overall our model and simulations did demonstrate how with minimal assumptions, novel variants can be introduced and spread in a population, leading to the eventual extinction of the existing variant. Other recent theoretical papers have aimed at modeling innovation as well as propagation of new linguistic forms. Blythe and Croft and Pierrehumbert et al. [10, 69] both presented models of the spread of a linguistic variant in a social network. Blythe and Croft
conclude that to create the S-curve typical of the propagation of a new linguistic variant, the innovator should be highly connected in a network, and the new variant should be associated with a group of high social status from an early stage (thus being assigned with a positive weighting that allows it to spread in a population). Pierrehumbert et al. presented a model where no positive weighting is assigned to the new variant, but its spread depends on general heterogeneous biases towards new variants among the speakers and their distribution in a network with variable connectivity. Both models represent the spread of a single innovation (introduced at one occasion by one speaker), thus accounting for language change in a neutral context where no pressure from language contact is involved. In these models, specific conditions, in terms of biases and/or network position of the innovator, are required for the novel variant to be successful. This implies a low probability of spread for any innovation occurring in the population, representing a good general match for non-contact induced spread. By contrast, our model assumes a homogeneously mixing population. This fact alone should not favor the spread of the novel variant. However, a more important difference between our model and previous work is the possibility of repeated introductions. This provides a sustained mutation pressure that, in the absence of directional forces (note that the model evolves neutrally), results in the eventual spread of the novel variant. The assumption that several L2 speakers introduce the same new variant independently of one another is grounded in second language acquisition theory, suggesting the occurrence of certain general effects in this process, such as for example morphological reductions, as well as influence from features from the speakers’ native language. The results of our simulations show that the multiple introductions of a novel variant allows a linguistic innovation to spread without any of the necessary conditions appointed for non-contact-induced change being fulfilled. We thus suggest that this may be a basic typological difference between contact-induced and non-contact-induced language change, which would explain how SLA may
accelerate language change in shift situations.
Figure 3.1: Change in the population mean usage of the mutant variant (averaged across ten runs) over 100 years. The other model parameter values are birth rate, $b=.007$, recruitment rate, $r=.063$, learning parameter, $l=.1$. For fixed population runs, $N=100$, while for expanding populations, $N_0 = 100$, with a growth rate of $g=b+r-d=.05$; $\mu$ is the mutation rate among $L2$ speakers. The trajectories of the expanding population means show more stability due to lower variance across runs.
Comparison between model and data

![Comparison between model and data](image)

**Figure 3.2:** Comparison of verb and preposition usage data with the mean model trajectories that most closely approached the data points for 1997, \((l=.05, \mu=.05), (l=.5, \mu=.05), (l=.7, \mu=.05), (l=.9, \mu=.05)\). Both data sets show greater spread than predicted by the model in the third growth phase (1997-2007), though the slope of the preposition data is reduced in qualitative agreement with the model. The verb data actually shows an increase in the spread of the novel variant, which conflicts with the model predictions.
Figure 3.3: Comparison of verb and preposition usage data with the model 2 mean trajectories that most closely matched the data points. The preposition use data fits the model predictions for $\mu = 0.0078$. The verb form data for 1997 fell within the 95% confidence intervals for the $\mu = 0.0063$ runs, but outside them for 2007.
Chapter 4

On measuring behavioral synchrony

Abstract

From breeding to flocking, synchrony is an important feature of many social behaviors. Different measures of synchrony have been proposed for different behaviors. Here we test how well some of these measures behave when applied to what we term timing and duration variable scalar behaviors (TDVS). These behaviors are those that may vary in timing (onset) and duration, and which can be characterized by a scalar variable at any time. Using agent-based simulations, we compare the effectiveness of four synchrony measures and show that two, the Dispersion Index of Mean Behavior, and the Kappa coefficient, perform best. We also show that population size affects the statistical interpretation of these measures. Finally, we also include additional results to show the relationship between a number of these measures.
4.1 Introduction

Synchrony is an important feature of many social behaviors. For example, synchrony in breeding activity affects operational sex ratios [15], while synchrony in other activities is crucial for the maintenance of social groups in space and time [49], particularly for the extraordinary flocking and swarming behaviors observed in some bird, fish, and insect species [17, 44, 82]. Before we can understand its evolutionary and ecological consequences we must first be able to measure the extent of synchrony. In the literature, a number of methods have been proposed for measuring the amount of synchrony in a population or group. For example, Kempenaer’s index [48] has been used for measuring breeding synchrony, while the average pairwise covariance in behavior [19] has been used in the collective behavior literature. However, much of the literature on behavioral synchrony has used idiosyncratic measures that lack formal justification and are often difficult to compare between studies [8, 23, 79]

The reason for the proliferation of synchrony measures stems from the inherent difficulty in measuring synchrony. Trying to determine to what extent animals are behaving in unison can require the acquisition of large amounts of data that are difficult to gather. Additionally, synchrony measures are not one size fits all, and must be chosen appropriately for the behavior in question, for instance, whether synchrony across all behaviors or only a particular behavior is being investigated. Additionally, without proper statistical reasoning attempts to measure synchrony can lead to spurious results.

One of the greatest difficulties in measuring synchrony is deciding which measure will perform best. An empirical studies rely on measures of synchrony to accurately detect the extent to which animals are behaving in unison. However, measures of synchrony are rarely tested in
controlled conditions to determine whether they do in fact identify synchrony when it is present. While it is difficult to manipulate study organisms in such a way to produce these controlled conditions, simulation models offer an opportunity to test the measures under known conditions. Asher Collins (2012) [5] used simulations to assess the performance of four synchrony measures, however their simulation models were limited and some of the measures lacked intuitive or formal justification. Here we study the behavior of four synchrony measures—Kempenaer’s Index, Simpson’s Index, the Kappa coefficient, and the dispersion index of mean behavior—under a number of different models and report the results.

Our models assume that animals can perform in what we term \textit{timing and duration variable scalar behaviors} (TDVS). TDVS behaviors are those for which the onset and the length of action may vary, and which can be described by a scalar measure (excluding vector-valued behaviors, such as velocity of a bird in a flock). Examples of TDVS behaviors are mating displays, breeding behavior, feeding activity, and vocalizations. We propose the use of the dispersion index of behavioral mean behavior as a simple method for testing for the presence of synchrony. We test this method against other measures on simulated data from different behavioral conditions. Additionally, we demonstrate the relationship between the dispersion index of mean behavior and the average pairwise covariance in behavioral state.

4.1.1 Synchrony Measures

We combed the literature for common synchrony measures, particularly for TDVS behaviors. While many studies use idiosyncratic measures that are unique to a particular study, we did identify some measures that have been used repeatedly: Kempenaer’s index for breeding synchrony, Simpson’s Index, and the Kappa coefficient.
4.1.1.1 Kempenaer Index

This measure was introduced to measure synchrony in the active breeding period (usually fixed in duration) which can occur at any time within some period [48, 60]. We begin with a focal individual, $i$, and calculate,

$$F_i = \frac{1}{t_i} \sum_{t \in t_i} \frac{n_i^t}{N-1} \quad (4.1)$$

where $t_i$ is the total time that $i$ behaved and $n_i^t$ is the number of individuals excluding $i$ that are displaying at time $t$ (the sum is over all the time steps where $i$ was behaving). 4.1 is the average proportion of the population that is performing the behavior of interest with $i$ at any time step. We then average this value over all individuals in the population,

$$F = \frac{1}{N} \sum_{i=1}^{N} F_i \quad (4.2)$$

$F$ takes values between 0 and 1, where 0 corresponds to no overlap in behavior, and 1 corresponds to 100% overlap.

4.1.1.2 Simpson Index

The Simpson Index, was introduced as a measure of diversity [76] and has been used to measure behavioral synchrony [49]. It is given by

$$SI_t = \sum_{i=1}^{b} \frac{n_{it}(n_{it} - 1)}{N(N-1)} \quad (4.3)$$

and is the probability that two individuals selected at random are engaging in the same behavior at time $t$, where $b$ is the number of behavioral categories. To measure synchrony we then average
these values over all time steps,

\[ SI = \frac{1}{T} \sum_{t=1}^{T} SI_t. \quad (4.4) \]

As \( SI \) is the average of a probability, it takes values from 0 to 1; \( SI = 0 \) means that no individual engages in the same behavior as another at the same time, while \( SI = 1 \) means that individuals always engage in the same behaviors at the same times.

### 4.1.1.3 Kappa Coefficient

The Kappa coefficient was introduced by Rook & Penning [73] to measure synchrony across behavior in sheep. For a given behavior at a certain time step, we calculate the proportion of pairs of animals (out of all possible pairings) that are engaging in that behavior. For a behavior \( i \), this is,

\[ p_t = \sum_{i=1}^{b} \frac{n_{it}(n_{it} - 1)}{N(N - 1)} \quad (4.5) \]

and is equivalent to 4.3 and thus is the probability that two randomly selected individuals at time \( t \) are engaging in behavior \( i \). The value \( P(A) \) is the average of 4.5 across all time steps, just as in 4.4,

\[ P(A) = \frac{1}{T} \sum_{t=1}^{T} p_t \quad (4.6) \]

Next we calculate \( P(E) \), the probability that two individuals chosen at random at randomly selected time steps (that is to say, two individuals selected at random from the entire data set) are engaging in the same behavior \( i \). As we are considering only two behaviors in this study (e.g. displaying or not displaying), let \( n_{it} \) be the number of individuals engaging in the behavior...
at time \( t \). Then \( P(E) \) is,

\[
P(E) = \sum_{i=1}^{b} \left( \frac{\sum_{t=1}^{T} n_{it}}{TN} \right) \left( \frac{\sum_{t=1}^{T} n_{it} - 1}{TN - 1} \right)
\]

(4.7)

Where again, \( b \) is the number of behavioral categories; in this paper \( b = 2 \) for all measures. The kappa coefficient is then given by,

\[
\kappa = \frac{P(A) - P(E)}{1 - P(E)}
\]

(4.8)

Thus, \( \kappa \) is how much more (or less) probable it is that two randomly selected individuals are engaging in the same behavior at the same time step than at different time steps, normalized by \( 1 - P(E) \), the probability that they are not engaging in the same behavior at different time steps. Like the other measures, \( \kappa \) takes values from 0 to 1, where 1 means there is total synchronization and 0 means there is no synchrony greater than that expected by chance. (Note: to be consistent with how we measure synchrony in our simulations, we use the explicit form of 4.7 which differs slightly from the large \( T \) approximation given in Rook & Penning (1993).)

4.1.1.4 Dispersion Index of mean behavior/Average pairwise covariance.

Based on our intuition that the synchrony should induce correlations between individuals’ behavior, we also considered the average pairwise covariance (APC) between individuals (a method which has also been used in studies of collective behavior for vector-valued behaviors [19]). However, calculating the APC requires knowing the behavior of each individual at all time steps, a difficult task in some studies when the animals may be moving around in the study area and difficult to distinguish from one another. Instead we calculated the dispersion index of mean
behavior (DIMB), which we show is directly related to the APC. By mean behavior we mean
the population mean of the behavioral states of all individuals at a time step. Let the behavioral
state of individual \( i \) at time \( t \) be given by \( \delta_{it} \); then the population mean behavior at time \( t \) is,

\[
\theta_t = \frac{1}{N} \sum_{i=1}^{N} \delta_{it}
\]  

(4.9)

Then the dispersion index of mean behavior is,

\[
D = \frac{\text{Var}(\theta)}{\theta}
\]  

(4.10)

Where \( \text{Var}(\theta) \) is the variance in \( \theta \). In Appendix G, we show that,

\[
D = \frac{E[Cov(\delta_{it}, \delta_{jt})]}{\theta}
\]  

(4.11)

where \( \text{Cov} \) is the covariance. Given the relationship with the APC, we can give the values of
the DIMB a simple interpretation. The DIMB takes values of 0 or greater, where \( \text{DIMB} = 0 \)
means that a fixed proportion of the population is engaged in a behavior at all time steps (no
synchrony).

**4.2 Methods**

The first difficulty in measuring synchrony is determining whether a given measurement dif-
fers significantly from what would be expected by chance. Marsden & Evans [60] recommend
measuring the synchrony over randomized replicates (>1000) of the data to generate a null
distribution of the measure and testing for significance of the true measurement. This means an
important distinction: a synchrony measure is really a statistic, and not a pure measurement. For instance, the mean height of a group of individuals sampled from a population is a statistic, while the measurement of a single individual is a pure measurement. The group mean allows for an estimate of the population mean, but any comparison must take into account chance variation due to the distribution of sample values. By contrast, any single measurement of an individual’s height can be compared to another (e.g. to say that individual 1 is taller than individual 2). Any synchrony value calculated from data collected on a group of animals is the result of a sample, and not a pure measurement. While this distinction may seem obvious, there are numerous instances of synchrony measures being interpreted as pure measurements and not statistics. Therefore, we follow the same procedure as Marsden & Evans in the rest of the paper.

Synchrony in action may result from a mix of factors. At the extremes, animals may respond only to social cues (e.g. the actions of neighbors), or be indifferent to their neighbors and react purely to common stimuli (e.g. the presence of a predator). In the former case, action is coordinated due to social interaction; in the latter case, the animals only have a shared response to the common stimulus. Additionally, synchrony may be observed even when animals are behaving randomly by pure chance, indifferent to either social or external stimuli. In order to determine the effectiveness of different measures of synchrony, we simulated populations behaving according to each of the three scenarios given above. By fixing the underlying behavioral rule, we can test whether a measure can distinguish scenarios in which animals should be expected to behave in synchrony from those where they should not.

In our first round of models—what we termed, “simple” models—animals calculated probabilities of behavior at each time step based on two ‘inputs’: the first was an individual motivation
level, given by \( p_0 := [0, 1] \); the second was an external signal, either the proportion of animals displaying at the previous time step, or an environmental signal. This gave us three model conditions:

1. **Uncoordinated Behavior**: Animals receive no input either from social or environmental cues and behave with fixed probability, \( p_0 \).

2. **Conformity Bias**: Animals behave with a probability that is a weighted average of their individual motivation (\( p_0 \)) and a probability determined by the proportion of the population that behaved in the previous time step,

   \[
   p_t = (1 - w_e)p_0 + w_e \left( \frac{1}{1 + \exp(-0.5)} \right)
   \]  

   where \( \pi \) is the proportion of the population that has displayed in the previous time step. Animals have a conformity bias, that is to say the rightmost term in 4.12 will increase when \( \pi > 50\% \) population has behaved and decrease when \( \pi < 50\% \) has behaved. The weight given to the external signal is \( w_e \), with \( (1 - w_e) \) being the weight given to the internal motivation. The greater \( w_e \) the more we should expect animals to act in unison.

3. **Shared Response**: Animals behave with a probability that is a weighted average of their individual motivation (\( p_0 \)) and a probability determined by an external signal from the environment,

   \[
   p_t = (1 - w_e)p_0 + w_e \left( \frac{1}{1 + \exp(-\lambda)} \right)
   \]  

   where \( \lambda \) represents an environmental cue and takes values between -1 and 1. The weight given to the external signal is \( w_e \), with \( (1 - w_e) \) being the weight given to the internal motivation. Again, the greater \( w_e \) the more we should expect animals to act in unison.
In the models above there is no bound on how often an animal may behave within the observation period. If the individual motivation and/or the external cue remains high, an animal may behave a majority of time steps. However, an individual’s behavior at any point in time will also be correlated with its previous behavior, and this can introduce noise in the signal of synchrony. In our next round of models—the complex models—we include this noisy element in our simulations by giving the animals a ‘mean behavioral bout’ time. When animals have been performing a behavior for some number of consecutive time steps less than this time, they are more likely to continue behaving. When they have been behaving for longer than this mean bout time, they reduce their probability of behaving. The probability of an animal engaging in a behavior is then the result of three inputs: 1) One’s own motivation; 2) the effect of one’s own behavioral history; 2) Whatever external cues influence behavior. The mathematical details of these simulations are given in Appendix I.

Synchrony is an outcome and not a cause of animal behavior. As such, our models only specify rules under which we should or should not expect synchrony to arise. The uncoordinated condition is the model least likely to produce synchronous behavior, as animals are behaving randomly and independent of one another (Note: this condition is equivalent to the two other conditions for $w_e = 0$). In the other two conditions, conformity bias and shared response, animals’ decisions to behave are coordinated by either social or environmental information and we expect them to induce synchrony. The shared response condition is most likely to induce synchrony, since the animals are able to respond simultaneously to the same stimulus.

We measured the behavioral synchrony across output from these simulations using the measures listed in the previous section. After each simulation we randomized the data to produce
a null distribution of measure values and tested the true value of the measure at the 5% significance level. We expected a useful synchrony measure to distinguish between the uncoordinated and the synchrony-inducing conditions (conformity bias and shared response), by producing a statistically significant value a majority of times. Additionally, we expect the rate at which statistically significant results appear to be greater for the shared response than for the conformity bias condition. Finally, within the two synchrony-inducing conditions, we expected the rate at which statistically significant levels of synchrony were detected to be correlated with values of $w_e$. Given that high individual motivation for behavior might result in synchrony purely from overproduction of behavior, we restricted the values of $p_0$ to $[0,.5]$. For each value of $p_0$ we ran simulations with $w_e = [.1,.2,.3,.4,.5,.6,.7,.8,.9]$. We report our results below.

4.3 Results

For each synchrony measurement we produced 1000 randomized replicates of the data to build a null distribution of values for the measure. We then tested for significance at the 5% level and reported significant measures as a ‘success’. The success rate ($S$) was the proportion of identically parametrized (some pair of values for $p_0$ and $w_e$) runs for which a ‘success’ was reported.

4.3.1 Simple Models

For $N = 10$, none of the measures reported $S > .5$ for the uncoordinated model condition, as expected. Only DIMB and Kappa resulted in $S > .5$ in the conformity bias and shared response conditions, meaning that there were parameter pairs under which it was more likely than not
that these measures would result in statistically significant levels of synchrony. As expected, there were more parameter pairings under the shared response condition with $S > .5$ than for the conformity bias condition (Detailed results from the simulations are reported in Appendices J and K).

At $N=50$, DIMB and Kappa resulted in $S > .5$ for the higher $p_0$ values ($> .3$) even in the uncoordinated condition, though the $s$ values remained higher for the synchrony-inducing conditions (and again higher for the shared response condition). The Simpson Index also showed $S > .5$ for the shared response condition. At $N = 100$ the same pattern continued with $D$ and $\kappa$ resulting in $S > .5$ for higher $p_0$ values, even in the uncoordinated condition, though still more so for the synchrony-inducing conditions. For the Simpson Index there were again values $S > .5$ for the shared response condition. The Kemenaer Index never reported statistically significant values for any of the model conditions.

The DIMB and Kappa measures most consistently reported significant synchrony values in the synchrony-inducing conditions, but they also reported higher $s$ values for the uncoordinated condition as the population size increased. We investigated this effect of population size on the rate at which statistically significant synchrony values were identified. We found that as the population size increased, the null distributions of the measures we produced by randomizing the data had decreased variance and lower mean. Essentially, as the population size increases, the null distributions were converging on a dirac function centered at the theoretical “no synchrony” value for these measures. Thus, any value greater than that theoretical minimum was significant at the 5% level (see Fig. 4.1). Another way to phrase this effect is that as the population size increased, these synchrony measures—which as we’ve already described are statistics—behaved
closer and closer to pure measurements. This is an important effect that should be considered in future studies.

### 4.3.2 Complex Models

The results for the complex models follow the exact same pattern as those for the simple models. The only difference was that overall, synchrony was detected less often for all measures than in the simple models. This is exactly what we’d expect, since introducing the correlation between an animal’s current and previous behavior should introduce noise into the system and make synchrony harder to detect.

**Table 4.1:** Correlation of success rate with \( w_e \) in simple models.

<table>
<thead>
<tr>
<th></th>
<th>Conformity bias</th>
<th></th>
<th>Shared response</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( F )</td>
<td>( SI )</td>
<td>( \kappa )</td>
</tr>
<tr>
<td>N=10</td>
<td>( p_0 = .1 ) ( \zeta^* ) ( \zeta^* ) ( .69^{<strong>} ) ( .72^{</strong>} ) ( \zeta^* ) ( \zeta^* ) ( .96^{<strong>} ) ( .97^{</strong>} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p_0 = .2 ) ( \zeta^* ) ( \zeta^* ) ( .72^{<strong>} ) ( .76^{</strong>} ) ( \zeta^* ) ( \zeta^* ) ( .89^{<strong>} ) ( .95^{</strong>} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p_0 = .3 ) ( \zeta^* ) ( \zeta^* ) ( .8^{<strong>} ) ( .8^{</strong>} ) ( \zeta^* ) ( \zeta^* ) ( .94^{<strong>} ) ( .94^{</strong>} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p_0 = .4 ) ( \zeta^* ) ( \zeta^* ) ( .91^{<strong>} ) ( .85^{</strong>} ) ( \zeta^* ) ( \zeta^* ) ( .95^{<strong>} ) ( .97^{</strong>} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p_0 = .5 ) ( \zeta^* ) ( \zeta^* ) ( .86^{<strong>} ) ( .87^{</strong>} ) ( \zeta^* ) ( \zeta^* ) ( .85^{<strong>} ) ( .82^{</strong>} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=50</td>
<td>( p_0 = .1 ) ( \zeta^* ) ( \zeta^* ) ( .87^{<strong>} ) ( .89^{</strong>} ) ( \zeta^* ) ( \zeta^* ) ( .85^{<strong>} ) ( .79^{</strong>} ) ( .79^{**} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p_0 = .2 ) ( \zeta^* ) ( \zeta^* ) ( .91^{<strong>} ) ( .9^{</strong>} ) ( \zeta^* ) ( \zeta^* ) ( .88^{<strong>} ) ( .81^{</strong>} ) ( .8^{**} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p_0 = .3 ) ( \zeta^* ) ( \zeta^* ) ( .91^{<strong>} ) ( .9^{</strong>} ) ( \zeta^* ) ( \zeta^* ) ( .89^{<strong>} ) ( .79^{</strong>} ) ( .8^{**} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p_0 = .4 ) ( \zeta^* ) ( \zeta^* ) ( .77^{<strong>} ) ( .7^{</strong>} ) ( \zeta^* ) ( \zeta^* ) ( .92^{<strong>} ) ( .68^{</strong>} ) ( .7^{**} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p_0 = .5 ) ( \zeta^* ) ( \zeta^* ) ( .8^{<strong>} ) ( .8^{</strong>} ) ( \zeta^* ) ( \zeta^* ) ( .93^{**} ) ( .52 ) ( .52 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=100</td>
<td>( p_0 = .1 ) ( \zeta^* ) ( .52 ) ( .97^{<strong>} ) ( .95^{</strong>} ) ( \zeta^* ) ( \zeta^* ) ( .89^{**} ) ( .61 ) ( .61 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p_0 = .2 ) ( \zeta^* ) ( .52 ) ( .87^{<strong>} ) ( .84^{</strong>} ) ( \zeta^* ) ( \zeta^* ) ( .87^{**} ) ( .67 ) ( .62 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p_0 = .3 ) ( \zeta^* ) ( \zeta^* ) ( .73^{<strong>} ) ( .73^{</strong>} ) ( \zeta^* ) ( \zeta^* ) ( .9^{<strong>} ) ( .52 ) ( .7^{</strong>} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p_0 = .4 ) ( \zeta^* ) ( .61 ) ( \eta^* ) ( \eta^* ) ( \zeta^* ) ( \zeta^* ) ( .86^{**} ) ( \eta^* ) ( \eta^* )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p_0 = .5 ) ( \zeta^* ) ( .8^{<strong>} ) ( \eta^* ) ( \eta^* ) ( \zeta^* ) ( \zeta^* ) ( .84^{</strong>} ) ( \eta^* ) ( \eta^* )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2: Correlations between success rate and \( w_e \) for all parameter pairs in the simple simulations, where success rate is defined as the proportion of times that a measure detected synchrony with \( p > .05 \). \( \zeta^* \) and \( \eta^* \) represent NA values due to the success rate being 0 and 1, respectively.
4.3.3 Synchrony-Inducing conditions

The greater the weight \( w_e \) given to an external signal, the greater the level of synchrony we expected. Table 4.2 shows the correlation between the success rate and \( w_e \) for all four measures under the two synchrony inducing conditions for the simple models. Overall, the DIMB and Kappa coefficient most consistently identified synchrony in the two conditions. However, as population size increased, the Simpson Index also began to show a significant correlation between the success rate and \( w_e \). As \( p_0 \) increased we also expected synchrony levels to rise, as animals were more likely to be engaging in the behavior and overlapping in time. Across measures we found more significant correlation values in the shared response than in the conformity bias condition, as expected.

Table 4.4 shows the correlations between success rate and \( w_e \) for the complex models that include the effect of a bout length. Here we see an even higher number of significant correlations, showing that the measures may be compensating for the overproduction of behavior in the simple models. Again, the DIMB and Kappa coefficient are the most consistent measures in producing a significant correlation between success rate and \( w_e \).

4.4 Discussion

To overcome the difficulty of determining how well synchrony measures behave in empirical studies, we used agent-based simulations where we could specify the underlying behavior of the agents. We then tested four measures of behavioral synchrony for a relevant class of behaviors (TDVS). We found that the most reliable measures were the DIMB and Kappa coefficients. These measures have the same data requirement conditions, and both have a simple intuitive
Table 4.3: Correlation of success rate with $w_c$ in complex models.

<table>
<thead>
<tr>
<th>Conformity bias</th>
<th>Shared response</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F$</td>
<td>$S1$</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>$p_0 = .1$</td>
<td>$\zeta^*$</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>$\zeta^*$</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>$\zeta^*$</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>$\zeta^*$</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>$\zeta^*$</td>
</tr>
</tbody>
</table>

Table 4.4: Correlations between success rate and $w_c$ for all parameter pairs in the complex simulations, where success rate is defined as the proportion of times that a measure detected synchrony with $p > .05$. $\zeta^*$ and $\eta^*$ represent NA values due to the success rate being 0 and 1, respectively.

interpretation, unlike a number of measures that have been used in the literature. Specifically, the DIMB is directly related to the average pairwise covariance in behavioral state, though it is both simpler with regard to data collection and calculation.

Asher & Collins [5] also used an agent-based simulation to assess the behavior of a number of synchrony measures, though we excluded most of the measures they included due to lack of intuitive and formal justification. Their study also found the Kappa coefficient to perform the best, though they did not include what we call the DIMB, or the average pairwise covariance.

Our results support the strong performance of the Kappa coefficient overall.

In conclusions, we recommend the use of the DIMB and Kappa coefficients as measures of behavioral synchrony for TDVS behaviors. These measures perform best across the multiple
conditions we simulated, and require less data collection. We showed that population size has an effect on these synchrony measures. As the population size increases, the measures behave more like pure measurements, which means that for studies of large populations these synchrony values can more justifiably be seen as measures of the level of synchrony, and direct comparisons can be made. Overall, increased consistency in the use synchrony measures will allow for better interpretation of experimental results and better comparison across empirical studies.
Appendix A

Expected time to MRUA

To solve equation (1.6), we first multiply through by $\kappa_k$ and rearrange the terms to obtain,

$$\mu_k (E[\tau_k] - E[\tau_{k-1}]) - \lambda_k (E[\tau_{k+1}] - E[\tau_k]) = 1 \quad (A.1)$$

By eq. 3.56 in Nåsell (2011) [62] we have,

$$E[\tau_k] - E[\tau_{k-1}] = \frac{1}{\mu_1 \rho_k} \sum_{j=k}^{N} \pi_j, \quad (A.2)$$

where we have made use of the second boundary condition in (1.7) and,

$$\rho_k = \frac{\lambda_1 \lambda_2 \ldots \lambda_{k-1}}{\mu_1 \mu_2 \ldots \mu_{k-1}} \quad (A.3)$$

$$\pi_j = \frac{\lambda_1 \lambda_2 \ldots \lambda_{j-1}}{\mu_2 \mu_3 \ldots \mu_j} \quad (A.4)$$
Imposing the boundary condition $E[\tau_1] = 0$ leads to,

$$E[\tau_k] = \frac{1}{\mu_1} \sum_{i=2}^{k} \sum_{j=1}^{N} \frac{\pi_j^i}{\rho_i}$$  \hspace{1cm} (A.5)

where, for convenience, we have moved the factor $1/\rho_j$ into the innermost sum. Lastly, writing $\rho_i$ and $\pi_j$ in terms of birth and death rates (equation 1.4) gives the solution in equation (1.8).

It is also easy to derive an expected time until sample size reaches a size $m$ smaller than our initial sample size but larger than one. If we define $\tau_{k,m} = \inf\{X_t = m; X_0 = n, n > m\}$ we have

$$E[\tau_{k,m}] = \frac{1}{r} \sum_{i=m+1}^{n} \sum_{j=1}^{N} \frac{1}{j} \left[ \frac{a}{r(N-1)} \right]^{j-1} \frac{N^j}{N^i},$$  \hspace{1cm} (A.6)

which we arrive at with the same argument as above, but with the revised boundary conditions

$$E[\tau_{m,m}] = 0, \quad E[\tau_{N,m}] = \frac{1}{\kappa_N} + E[\tau_{N-1}].$$  \hspace{1cm} (A.7)

### A.1 Variance in time to MRUA

We begin with the more general form of (1.6) [63],

$$E_i^r = \frac{r}{\lambda_i + \mu_i} E_i^r + \frac{\lambda_i}{\lambda_i + \mu_i} E_{i+1}^r + \frac{\mu_i}{\lambda_i + \mu_i} E_{i-1}^r$$  \hspace{1cm} (A.8)

where $E_i^r = E(\tau_i^r)$, the $r$th moment of the random variable $\tau_i$. For $r = 2$ we have boundary conditions,

$$E_1^2 = 0, \quad E_1^1 = 0, \quad E_N^1 = \frac{1}{\lambda_N + \mu_N} + E_{N-1}^1$$  \hspace{1cm} (A.9)
Using (A.9), we can derive the following equation for the second moment (see eq. 6.7, [? ]),

\[ E_i^2 = 2 \sum_{j=2}^{N} E_j^{1} \alpha_j \sum_{k=2}^{\min(i,j)} \psi_k \] (A.10)

where \( \alpha_j = (\mu_j \psi_j)^{-1} \) and \( \psi_k = \frac{\lambda_k \lambda_{k+1} \cdots \lambda_N}{\mu_k \mu_{k+1} \cdots \mu_N} \). Reversing the order of the sums and moving all terms to the innermost sum yields,

\[ E_i^2 = 2 \sum_{k=2}^{i} \sum_{j=2}^{N} E_j^{1} \alpha_j \psi_k \] (A.11)

Substituting the definitions for \( \alpha_j \) and \( \psi_k \), we have the state dependent second moments,

\[ E_i^2 = 2 \sum_{k=2}^{i} \sum_{j=2}^{N} E_j^{1} \frac{1}{r(N-1)} \left[ \frac{a}{r(N-1)} \right]^{j-k} \frac{N^k}{N^k}. \] (A.12)
Appendix B

Expected time to MRUA in

Strimling et al. model

Common ancestry events occur when a replacement event has occurred within the sample and the new individual has chosen another sample member as her cultural parent. Replacement events occur with rate $rN$, where $r$ is the reciprocal of average lifetime and $N$ is the population size. Given that a replacement event has occurred, the probability that replacement and learning both take place within the sample is $j(j - 1)/[N(N - 1)]$. Therefore, a common ancestry event will occur in a sample of size $j$ with rate,

$$\frac{rj(j - 1)}{N - 1} \quad \text{(B.1)}$$
and the expected time to this event will be \((N - 1)/[r(j - 1)]\). Note that \(\tau_n\), the time to MRUA for an initial sample \(X_0 = n\), is

\[
\tau_n = t_1 + t_2 + \cdots + t_{n-1},
\]

(B.2)

where \(t_i\) is the time between the \(i\)th and \((i - 1)\)th common ancestry event. Thus, \(E[\tau_n] = E[t_1] + E[t_2] + \cdots + E[t_{n-1}]\), and,

\[
E[\tau_n] = \frac{N - 1}{rn(n - 1)} + \frac{N - 1}{r(n - 1)((n - 1) - 1)} + \cdots + \frac{N - 1}{2r} = \sum_{j=2}^{n} \frac{(N - 1)}{r(j - 1)}
\]

(B.3)

For \(r = 1\) we have the result given in (1.12).
Appendix C

Algorithm for simulating the genealogical process

Below is the algorithm we used for simulating the genealogical process:

1. Let $N =$ population size, $r =$ replacement rate, $a =$ learning rate.

2. Let $k =$ initial sample size.

3. Let $c = 0$ (event counter).

4. Increment $c$ by 1.

5. Let $\lambda = a \times k \times (N - k)/(N - 1)$.

6. Let $\mu = r \times k$.

7. Let $x = \lambda / (\lambda + \mu)$.

8. Let $y =$ uniform random deviate in $[0,1]$. 
9. if \( y \leq x \), increment \( k \) by 1, else decrement \( k \) by 1.

10. If \( k \) is 1, print out \( c \) and stop, otherwise go to step 4.

This algorithm simulates events. If one wants to simulate time one can increment \( c \) by an exponential deviate with expected value of \( (r + a) \times N \), then \( c \) will measure time.
Appendix D

System Measures and Cultural Systems Simulation Model details

D.0.1 Compatibility calculations

The compatibility relationship between traits is one of the most important features of the model. In addition to influencing cultural transmission through the various filters, compatibility also offers a way to examine the macro scale effects of our model. We use three compatibility measures to explore the behavior of the model: 1) Internal compatibility measures the average compatibility of an individual's own traits; 2) Between individual compatibility measures the average compatibility among all pairs of individuals in the population; 3) Pool compatibility measures the overall compatibility of all traits in the population. Below we offer detailed descriptions of the calculation of each compatibility measure.
D.0.1.1 Internal Compatibility

In order to measure the level of compatibility among an individual’s traits, we calculate the average compatibility among all trait pairs, excluding self-comparisons. Thus, we calculate,

$$\text{internal compatibility} = \frac{\sum_{(i,j):i\neq j} R(i,j)}{|L/(i,j : i = j)|}$$  \hspace{1cm} (D.1)

D.0.1.2 Between individual compatibility

To find the average compatibility between individuals, we calculate the average compatibility across all trait pairs \((i, j)\), such that \(i \in I, j \in J\) where \(I\) and \(J\) represent the sets of traits of two distinct individuals. If \(|I|\) or \(|J| = 0\), we do not include the comparison in the average. If we define \(H := (I \times J)/(i, j) : i = j\), we have,

$$\text{between individual compatibility} = \frac{\sum_{(i,j):i\neq j} R(i,j)}{|H|}$$  \hspace{1cm} (D.2)

D.0.1.3 Pool compatibility

To characterize the pool of cultural traits in the population (as opposed to the universe of possible traits), we include a measure of the overall compatibility of the culture. Pool compatibility is measured by averaging over the compatibility values of all pairs of traits present in at least one individual, weighted by the occurrence of each trait. More precisely, if \(S_i\) represents the trait set of individual \(i\), then

$$\text{pool compatibility} = \frac{\sum_{(t_1,t_2):t_1\neq t_2} R(t_1,t_2)}{\prod_{i \in N} |S_i|}$$  \hspace{1cm} (D.3)
for $t_1 \in S_i, t_2 \in S_j$.

### D.0.2 Other system measures

The following list describes the measures of the systems reported in the results:

1. **Culture size**: the number of traits possessed by at least one individual in the population.

2. **Repertoire size**: the number of traits possessed by a single individual.

3. **Similarity**: The proportion an agent’s own traits that it shares with another agent. As the number of shared traits is normalized by the number of an agent’s own traits, similarity is not necessarily symmetric. For example, given agents 1 and 2, with repertoires $R_1$ and $R_2$, the similarity of agent 2 to agent 1 is,

$$\frac{|R_1 \cap R_2|}{|R_1|}$$

while the similarity of agent 1 to agent 2 is,

$$\frac{|R_1 \cap R_2|}{|R_2|}$$

### D.0.3 Simulation details

In calculating the probabilities of copying or innovation produced by the filters, we made use of a quantity, $s$, called the score. The exact calculation of the score depended on the filter in use, but the quantity was a way of encapsulating the information about compatibility relationships laid out in the trait universe, and so $-1 \geq s \geq 1$. 
The probability of copying under the TF and MF, and the probability of innovating under the IF, were calculated using the following logistic function:

\[ p = \frac{1}{1 + e^{-ks}} \]  \hspace{1cm} (D.4)

where the parameter \( k \) determines the strength of the filter effect. For the simulation results reported in the paper, \( k = 1 \), though varying \( k \) did not qualitatively change the results.

### D.0.3.1 Trait Filter

When an agent encounters a model it selects one of her traits, \( t \), at random for potential copying. Let \( L \) be the learner’s set of traits. We define the score, \( s \), as,

\[ s = \frac{\sum_{l \in L} R(t, l)}{|L|} \]  \hspace{1cm} (D.5)

where the sum is over all \( l \in L \). Copying occurs with a probability \( p \), given by D.4. Thus, the more compatible the trait \( t \) is on average with the learner’s traits, the more likely it will be copied.

### Model Filter

When an agent encounters a model it selects one of her traits, \( t \), at random for potential copying. Let \( L \) be the set of the learner’s traits and \( M \) the set of the model’s traits with \( l \in L \) and \( m \in M \). We now define the score, \( s \), as
where the sum is over all \((m, l) \in |M \times L|\). The probability of copying is given by D.4. Here, the learner considers the overall compatibility between herself and the model in determining the probability of copying.

### Innovation Filter

When an agent is given the opportunity to invent, they will select a trait, \(t\), at random from the trait universe, for potential innovation. The score is then calculated according to D.5, and the probability of copying by D.4. Here, individuals are more likely to innovate (i.e. introduce a trait into the population) a trait if it is on average compatible with their current repertoire of traits.

### Display Filter

When a learner selects a cultural model, the model selects a trait to make visible to the learner. Let \(M\) be the set of the model’s traits with \(m, n \in M\). The trait-specific score \(s_m\) is,

\[
s_m = \frac{\sum R(m, n)}{|M|} \tag{D.7}
\]

We then rescale the score as \(\frac{s_m + 1}{2}\) to values between 0 and 1. We rank the traits based on these rescaled values and assign them positions on the unit interval, \(x_m\), equal to their rescaled scores added to the position of the previous trait in the ranking. Next we select a random number, \(r\), on the unit interval; if \(x_{m-1} \leq r < x_m\), then the model will choose to display trait \(m\) and
only this trait will be available for copying in that encounter. Thus, we have used the average compatibility value of each trait in the repertoire to construct a distribution of probabilities of display across all a model’s traits.

**Simulation details**

The simulation runs included only one active filter at a time. Runs with pairs or triads of filters show that the effects were additive. The simulations were run with a population of 100 agents with average lifespans of 100 interactions. The simulation was run for $1e4$ rounds of interaction, or 100 model generations. All results are averaged over the final 20% of model runs, with 10 runs for each constellation of parameters.
Appendix E

Expected change in the mean frequency of the existing variant

In the model described in the text, for a fixed population of size $N$, we have some proportion $\rho N$ every time step who are L2 speakers. Each of these can ‘mutate’ (i.e. assign probability 1 to a variant other than the existing variant) with probability $\mu$. Thus, every time step there will be $\mu \rho N$ mutants in the population.

In each round of interaction every individual will be assigned to an exclusive pair. Let us call the set of all such possible assignments for the population $S$. For a given $s \in S$, we will have three possible types of pairings $(i, j)$:

1. Two non-mutants, whose probabilities of using the existing variant are $p_{i1}$ and $p_{j1}$.

2. A non-mutant and a mutant, whose probabilities of using the existing variant are $p_{i1}$ and 0, respectively.
3. Two mutants, whose probabilities of using the existing variant are both 0.

The change in the population mean frequency of the existing variant is equal to the sum of the changes in the probabilities of usage of that variant across the population; in other words, $\Delta \bar{p}_1 = \sum_{i=1}^{N} \Delta p_{i1}$, where $\Delta p_{i1} = p'_{i1} - p_{i1}$. Thus,

$$E[\bar{p}_1] = \sum_{i=1}^{N} E[\Delta p_{i1}] = \sum_{(i,j) \in s} E[\Delta p_{i1} + \Delta p_{j1}],$$  \hspace{1cm} (E.1)

since every individual is assigned to an exclusive pair. For a pairing of type 1, we have,

$$E[\Delta p_{i1} + \Delta p_{j1}] = E[\Delta p_{i1}] + E[\Delta p_{j1}]$$  \hspace{1cm} (E.2)

$$= (1 - p_{i1})lp_{j1} - p_{i1}l(1 - p_{j1}) + (1 - p_{j1})lp_{i1} - p_{j1}l(1 - p_{i1}) = 0$$  \hspace{1cm} (E.3)

For a pairing of type 2 (assuming $i$ is the non-mutant), we have,

$$E[\Delta p_{i1} + \Delta p_{j1}] = E[\Delta p_{i1}] + E[\Delta p_{j1}]$$  \hspace{1cm} (E.4)

$$= -lp_{i1} + lp_{i1} = 0$$  \hspace{1cm} (E.5)

And finally, for a pairing of type 3,

$$E[\Delta p_{i1} + \Delta p_{j1}] = 0$$  \hspace{1cm} (E.6)

No matter which $s \in S$ we choose, the sum over all pairings is 0; thus $E[\bar{p}_1] = 0$, and the process is neutral with respect to the existing variant.
Appendix F

Maputo Data

We obtained demographic data on the number of $L1$ and $L2$ speakers of Portuguese in Maputo, Mozambique in the years 1975, 1980, 1997, and 2007 (see Table 3.1). This period is important because it captures the dramatic expansion of Portuguese speakers in Maputo following the official independence of the Republic of Mozambique in 1975, during which Mozambiquan Portuguese has undergone a number of changes. We assumed geometric growth between each time point and estimated the rates at which $L1$ and $L2$ speakers entered the population ($b$ and $r$) in the following way:

1. Fix the overall growth rate ($g$) of the Portuguese speaking population based on the values in Table 3.1.

2. For each period (1975-1980, 1980-1997, 1997-2007), take a geometric average of the five year smoothed country wide birth ($b$) and death ($d$) rates for Mozambique.

3. Calculate the recruitment rate, $r = g + d - b$. 

95
The above method gave us demographic projections that matched the data very closely, and matched the proportions of $L1$ and $L2$ speakers at each time point particularly closely. We used these parameter estimates to simulate language change in Maputo according to the model describe above.

For this study we use empirical data concerning two variable grammatical phenomena: reduced verbal agreement and innovative use of prepositions:

1. **Verbal Agreement**: For reduced verbal agreement, we include verbs where the subject is a plural pronoun or a plural noun phrase. The standard variant (assumed to be used by all speakers at the time of independence) is the verb exposing agreement with the plural subject by a plural suffix. The new variant is the lack of the plural suffix on the verb, the verb assuming the morphologically simpler singular form.

2. **Preposition Use**: Innovative use of prepositions that implies excluding, adding or changing any of the prepositions *de* (of), *a* (to), *em* (in), *por* (for), *com* (with), *para* (to), or *sobre* (on). The existing variant is the use of the a conservative preposition (as in European Portuguese) following verbs identified as exposing variable behavior in Mozambican Portuguese, in terms of which preposition they trigger. The new variant is the use of a new preposition (different from European Portuguese) following the identified verbs.

Reduced verbal agreement and innovative preposition use are features that have been appointed as typical results of language contact in Maputo Portuguese, other varieties of African Portuguese as well as other contact varieties of Portuguese [6, 59, 66]. Furthermore, these features occur in our datasets with a frequency high enough for us to distinguish a pattern for their
Appendix F. *Maputo Data* 

development. For both phenomena occurrences of new linguistic variants are compared to occurrences of standard variants.

The datasets come from recordings that were made in Maputo in 1997 and 2007 within the frameworks of studies by Stroud and Gonçalves (1997) and Jon-And (2011) [1, 81] respectively. Both samples consist in 20 recorded informal semi-structured interviews of 30-60 minutes with 20 participants. The participants in the two samples are comparable in terms of age groups and education levels, factors that have often been found to determine linguistic variation. The recordings are also made in similar circumstances and the interviews concern the same subject matters. For the earlier sample we have access only to quantitative results, presented by Stroud and Gonçalves in “Panorama do Português Oral de Maputo”, and not to the original recordings or transcriptions. For the later sample we have full access to the recorded and transcribed corpus. In an extensive comparison of the occurrences of new morphological, syntactic and lexical variants of all 20 features that are registered by Stroud and Gonçalves, the most common new forms are the same in the two samples, and similar tendencies are registered when comparing the relative frequency of new forms at the morphological, syntactic and lexical level, indicating a high level of reliability and comparability of the datasets (Reite and Jon-And, unpublished data). The access to these two datasets provides us with a unique opportunity for a quantitative chronological study of change in progress in ongoing language contact.

<table>
<thead>
<tr>
<th>period</th>
<th>recruitment rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975-1980</td>
<td>0.7590218</td>
</tr>
<tr>
<td>1980-1997</td>
<td>0.04677117</td>
</tr>
<tr>
<td>1997-2007</td>
<td>0.0128031</td>
</tr>
</tbody>
</table>

Table F.1: Estimated rate of recruitment of L2 speakers for Maputo Portuguese-speaking population.
Appendix G

Relationship between dispersion and covariance

Here we show how the time dispersion index relates to the average pairwise covariance. Let $\delta_{it}$ represent the behavioral state of individual $i$ at time $t$. Assume in the current example that the behavior is binary, such that, $\delta_{it} = [0, 1]$ (Note: we will generalize to all scalar behaviors below). The polarization—the average population behavior at time, $t$—is $\theta_t = \frac{1}{N} \sum_{i=1}^{N} \delta_{it}$.

The dispersion index is ratio of the variance in polarization to the mean polarization,

$$D = \frac{Var(\theta_t)}{\theta_t}$$  \hspace{1cm} (G.1)

The variance in polarization is,

$$Var(\theta_t) = E[\theta_t^2] - (E[\theta_t])^2.$$  \hspace{1cm} (G.2)
Appendix E. Relationship between dispersion and covariance

Now, substitute the definition of \( \theta_t \),

\[
\frac{1}{T} \sum_{t=1}^{T} \left( \frac{1}{N} \sum_{i=1}^{N} \delta_{it} \right)^2 - \left( \frac{1}{T} \sum_{t=1}^{T} \frac{1}{N} \sum_{i=1}^{N} \delta_{it} \right)^2 = \frac{1}{TN^2} \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{j=1}^{N} \delta_{it} \delta_{jt} - \frac{1}{(TN)^2} \sum_{t=1}^{T} \sum_{\tau=1}^{T} \sum_{i=1}^{N} \sum_{j=1}^{N} \delta_{it} \delta_{j\tau} \tag{G.3}
\]

Switching the order of the sums, and associating the appropriate factors, we have,

\[
Var(\theta_t) = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{1}{T} \sum_{t=1}^{T} \delta_{it} \delta_{jt} = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \left[ E[\delta_{it} \delta_{jt}] - E[\delta_{it}] E[\delta_{jt}] \right] \tag{G.5}
\]

\[
= E[Cov(\delta_{it}, \delta_{jt})] \tag{G.6}
\]

Thus, \( D = E[Cov(\delta_{it}, \delta_{jt})] / E[\theta_t] \). In demonstrating this relationship, we made no use of the fact that \( \delta_{it} \) is a a binary variable.
Appendix H

Synchrony simulation model details

We implemented two types of models to test different measures of synchrony. Below we describe the simulation models in detail.

H.0.1 Simple models

At each model time step an animal performs the behavior with a probability $p_i$. Each animal’s probability of behavior is a weighted average of their own motivation and some external input. The self motivation is a fixed baseline probability of behavior, $p_0$; the external input varies according to the simulation type.
H.0.1.1 Uncoordinated behavior

In this model animals behave with no external input. The probability of an individual $i$ engaging in the behavior of interest at a time step is,

$$p_i = p_0 \quad (H.1)$$

H.0.1.2 Conformist behavior

In this model animals are given a conformity bias: when a majority of the population is behaving, they are positively biased towards behaving; when a minority of the population are behaving, they are negatively biased towards behaving. The probability that an individual $i$ engages in the behavior is determined by the following equation,

$$p_i = (1 - w_e)p_0 + w_e \left( \frac{1}{1 + e^{-(\pi - .5)/.5}} \right) \quad (H.2)$$

Here, $w_e$ is the weight given to the external input and $\pi$ is the the fraction of the population that has displayed in the previous time step. The second term on the right in eq. 1.2 contains a logistic function that maps the conformity value, $(\pi - .5)/.5$—which takes values between -1 and 1—to a probability.

H.0.1.3 Shared response

In this model animals receive an environmental signal, $\zeta := [-1, 1]$. Positive values of $\zeta$ encourage behaving while negative values have the opposite effect. The probability of behavior for $i$
Appendix H. *Synchrony simulation model details*  

is given by the following equation,

\[ p_i = (1 - w_e)p_0 + w_e \left( \frac{1}{1 + e^{-\zeta}} \right) \]  \hspace{1cm} (H.3)

**H.0.2 Complex Models**

In these models we incorporate the idea that once an animal has commenced behavior, it is often likely to continue to do so for some period. We assign a mean bout time. If an animal has been engaging in the behavior for some number of consecutive time steps less than half the mean bout time, they increase their probability of behaving in the current time step. Once they have been engaging in the behavior for more than half the mean bout time, they decrease the probability until it reaches zero at the mean bout time. This introducing a kind of ‘gas pedal’ on the behavior; once an animal starts behaving it accelerates the gas so that it is more likely to behave again; after a point, the animal presses on the brake, and decelerates, so to speak. Details are given below.

**Uncoordinated condition**

Each animal begins with a baseline probability of behaving, \( p_0 \). At each time step, every animal reviews its previous behavior. Let \( s \) be the number of consecutive time steps spent engaging in the behavior thus far, and \( T \) be the mean bout length for an animal in the population.

If \( s < T/2 \), then the animal updates its probability of displaying in the current time step by:

\[ p'_i = p_i + \frac{s}{T/2} (1 - p_i) \]  \hspace{1cm} (H.4)
Appendix H. Synchrony simulation model details

Where $p_i$ and $p'_i$ are the probabilities of behaving in the previous and current time steps, respectively.

If $s \geq T/2$, the animal updates its probability of display by:

$$p'_i = p_i - \frac{s}{T}(p_i - p_0) \quad (H.5)$$

Thus the animal increases its probability of behaving when it has behaved for less than half of the mean bout time, and decreases its probability of behaving when it has behaved for more than half of the mean bout time. This introduces a correlation between an animal’s current and past behavioral states.

**Conformity bias condition**

In addition to the basic behavioral rule above, we now add a conformity bias. When $s = 0$, the animal updates its probability of behaving just as in the simple model,

$$p_i = (1 - w_e)p_0 + w_e \left( \frac{1}{1 + e^{-(\pi - .5)/.5}} \right) \quad (H.6)$$

However, if $0 < s$, then we have,

$$p'_i = p_i + (1 - w_e) \left[ \frac{s}{m} (1 - p_i)\delta_{0<s\leq T/2} - \frac{s}{T}p_i\delta_{T/2<s} \right] + w_e \left[ \left( \frac{\pi - .5}{.5} \right) p_i\delta_{0<\pi\leq .5} + \left( \frac{\pi - .5}{.5} \right) (1 - p_i)\delta_{.5<\pi} \right] \quad (H.7)$$

Where the $\delta_{i}$ are indicator functions taking the value 1 when the condition in the subscript is met and zero otherwise. The terms with the indicator functions ‘switch on’ when the conditions...
are met and ensure that the contributions of the terms increase or decrease appropriately, while ensuring that the overall probability remains between 0 and 1. As an example, let’s consider the probability that $i$ will engage in the behavior when it has been doing so for $s > T/2$ and $\pi = .3$,

$$p'_i = p_i - \left[ (1 - w_e) \frac{s}{T} + 0.4w_e \right] p_i \quad (H.8)$$

The rightmost term reduces the probability of engaging in the behavior, as it should, since the animal has been behaving for a while and a minority of the population has engaged in the behavior in the previous time step.

**Shared Response**

The shared response works in the same way as the conformity bias condition with only a different external input. When $s = 0$,

$$p_i = (1 - w_e)p_0 + w_e \left( \frac{1}{1 + e^{-\zeta}} \right) \quad (H.9)$$

just as above. However, if $0 < s$, then we have,

$$p'_i = p_i + (1 - w_e) \left[ \frac{s}{m} (1 - p_i)\delta_{0 < \zeta \leq m} - \frac{s}{2m} p_i \delta_{m < s} \right] + w_e \left[ \zeta p_i \delta_{\zeta \leq 0} + \zeta (1 - p_i) \delta_{\zeta > 0} \right] \quad (H.10)$$

The indicator functions $\delta$ work in the same way as for the conformity bias.
Appendix I

Synchrony simulation model details

We implemented two types of models to test different measures of synchrony. Below we describe the simulation models in detail.

I.0.1 Simple models

At each model time step an animal performs the behavior with a probability $p_i$. Each animal’s probability of behavior is a weighted average of their own motivation and some external input. The self motivation is a fixed baseline probability of behavior, $p_0$; the external input varies according to the simulation type.
I.0.1.1 Uncoordinated behavior

In this model animals behave with no external input. The probability of an individual $i$ engaging in the behavior of interest at a time step is,

$$p_i = p_0$$  \hspace{1cm} (I.1)

I.0.1.2 Conformist behavior

In this model animals are given a conformity bias: when a majority of the population is behaving, they are positively biased towards behaving; when a minority of the population are behaving, they are negatively biased towards behaving. The probability that an individual $i$ engages in the behavior is determined by the following equation,

$$p_i = (1 - w_e)p_0 + w_e \left( \frac{1}{1 + e^{-(\pi - .5)/.5}} \right)$$  \hspace{1cm} (I.2)

Here, $w_e$ is the weight given to the external input and $\pi$ is the the fraction of the population that has displayed in the previous time step. The second term on the right in eq. I.2 contains a logistic function that maps the conformity value, $(\pi - .5)/.5$—which takes values between -1 and 1—to a probability.

I.0.1.3 Shared response

In this model animals receive an environmental signal, $\zeta := [-1, 1]$. Positive values of $\zeta$ encourage behaving while negative values have the opposite effect. The probability of behavior for $i$
Appendix H. *Synchrony simulation model details*

is given by the following equation,

\[ p_i = (1 - w_e)p_0 + w_e \left( \frac{1}{1 + e^{-\zeta}} \right) \]  
(I.3)

I.0.2 Complex Models

In these models we incorporate the idea that once an animal has commenced behavior, it is often likely to continue to do so for some period. We assign a mean bout time. If an animal has been engaging in the behavior for some number of consecutive time steps less than half the mean bout time, they increase their probability of behaving in the current time step. Once they have been engaging in the behavior for more than half the mean bout time, they decrease the probability until it reaches zero at the mean bout time. This introducing a kind of ‘gas pedal’ on the behavior; once an animal starts behaving it accelerates the gas so that it is more likely to behave again; after a point, the animal presses on the brake, and decelerates, so to speak. Details are given below.

Uncoordinated condition

Each animal begins with a baseline probability of behaving, \( p_0 \). At each time step, every animal reviews its previous behavior. Let \( s \) be the number of consecutive time steps spent engaging in the behavior thus far, and \( T \) be the mean bout length for an animal in the population.

If \( s < T/2 \), then the animal updates its probability of displaying in the current time step by:

\[ p_i' = p_i + \frac{s}{T/2}(1 - p_i) \]  
(I.4)
Appendix H. Synchrony simulation model details

Where \( p_i \) and \( p_i' \) are the probabilities of behaving in the previous and current time steps, respectively.

If \( s \geq T/2 \), the animal updates its probability of display by:

\[
p_i' = p_i - \frac{s}{T} (p_i - p_0)
\]  

(1.5)

Thus the animal increases its probability of behaving when if it has behaved for less than half of the mean bout time, and decreases its probability of behaving when its has behaved for more than half of the mean bout time. This introduces a correlation between an animal’s current and past behavioral states.

Conformity bias condition

In addition to the basic behavioral rule above, we now add a conformity bias. When \( s = 0 \), the animal updates its probability of behaving just as in the simple model,

\[
p_i = (1 - w_e)p_0 + w_e \left( \frac{1}{1 + e^{-(\pi - .5)/.5}} \right)
\]

(1.6)

However, if \( 0 < s \), then we have,

\[
p_i' = p_i + (1 - w_e) \left[ \frac{s}{m} (1 - p_i) \delta_{0 < s \leq T/2} - \frac{s}{T} p_i \delta_{T/2 < s} \right] + w_e \left[ \left( \frac{\pi - .5}{.5} \right) p_i \delta_{0 < \pi \leq .5} + \left( \frac{\pi - .5}{.5} \right) (1 - p_i) \delta_{.5 < \pi} \right]
\]

(1.7)

Where the \( \delta \) are indicator functions taking the value 1 when the condition in the subscript is met and zero otherwise. The terms with the indicator functions ‘switch on’ when the conditions
are met and ensure that the contributions of the terms increase or decrease appropriately, while ensuring that the overall probability remains between 0 and 1. As an example, let’s consider the probability that \( i \) will engage in the behavior when it has been doing so for \( s > T/2 \) and \( \pi = .3 \),

\[
p_i' = p_i - \left[ (1 - w_e) \frac{s}{T} + 0.4w_e \right] p_i
\]

(I.8)

The rightmost term reduces the probability of engaging in the behavior, as it should, since the animal has been behaving for a while and a minority of the population has engaged in the behavior in the previous time step.

**Shared Response**

The shared response works in the same way as the conformity bias condition with only a different external input. When \( s = 0 \),

\[
p_i = (1 - w_e)p_0 + we \left( \frac{1}{1 + e^{-\zeta}} \right)
\]

(I.9)

just as above. However, if \( 0 < s \), then we have,

\[
p_i' = p_i + (1 - w_e) \left[ \frac{s}{m} (1 - p_i) \delta_{0 < s < m} - \frac{s}{2m} p_i \delta_{m < s} \right] + w_e \left[ \zeta p_i \delta_{\zeta \leq 0} + \zeta (1 - p_i) \delta_{\zeta > 0} \right]
\]

(I.10)

The indicator functions \( \delta \) work in the same way as for the conformity bias.
Appendix J

Detailed Simple model results

Simple model results
Appendix X. *Detailed Simple model results*

Table J.1: Dispersion Index of Polarization, N=10

<table>
<thead>
<tr>
<th>(w_e)</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>(p_0 = .1)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(p_0 = .2)</td>
<td>0</td>
<td>0.1</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>(p_0 = .3)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>(p_0 = .4)</td>
<td>0.1</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>(p_0 = .5)</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Table J.2: The proportion of runs for which the Dispersion Index of Polarization detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 10.

Table J.1: Dispersion Index of Polarization, N=10

<table>
<thead>
<tr>
<th>(w_e)</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>(p_0 = .1)</td>
<td>0</td>
<td>0.2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.6</td>
</tr>
<tr>
<td>(p_0 = .2)</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0.1</td>
<td>0</td>
<td>0.2</td>
<td>0</td>
<td>0.3</td>
<td>0.2</td>
<td>0.5</td>
</tr>
<tr>
<td>(p_0 = .3)</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
<td>0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.5</td>
<td>0.6</td>
</tr>
<tr>
<td>(p_0 = .4)</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.3</td>
<td>0.2</td>
<td>0.3</td>
<td>0.2</td>
<td>0.5</td>
<td>0.5</td>
<td>0.4</td>
</tr>
<tr>
<td>(p_0 = .5)</td>
<td>0.1</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
<td>0.5</td>
<td>0.8</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Table J.2: The proportion of runs for which the Dispersion Index of Polarization detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 10.
### Table J.3: Simpson Index, N=10

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### Table J.4: The proportion of runs for which the Simpson Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 10.

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Table J.5: Kempenaer Index, N=10

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table J.6: The proportion of runs for which the Kempenaer Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 10.

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
### Table J.7: Kappa coefficient, N=10

<table>
<thead>
<tr>
<th>$w_e$ =</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0.2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0.3</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0.1</td>
<td>0</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
</tr>
</tbody>
</table>

### Table J.8: The proportion of runs for which the Kappa coefficient detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 10.

<table>
<thead>
<tr>
<th>$w_e$ =</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0.2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0.2</td>
<td>0</td>
<td>0.3</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
<td>0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.3</td>
<td>0.1</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.3</td>
<td>0.2</td>
<td>0.5</td>
<td>0.5</td>
<td>0.4</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0.4</td>
<td>0.3</td>
<td>0.2</td>
<td>0.3</td>
<td>0.5</td>
<td>0.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$w_e$ =</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0.3</td>
<td>0.2</td>
<td>0.5</td>
<td>0.7</td>
<td>0.8</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
<td>0.4</td>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0.1</td>
<td>0.4</td>
<td>0.3</td>
<td>0.4</td>
<td>0.8</td>
<td>0.8</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.5</td>
<td>0.6</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.3</td>
<td>0.1</td>
<td>0.7</td>
<td>0.8</td>
<td>0.8</td>
<td>1.0</td>
</tr>
</tbody>
</table>
Table J.9: Dispersion Index of Polarization, N=50

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0.2</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0.3</td>
<td>0.1</td>
<td>0.3</td>
<td>0.4</td>
<td>0.3</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.3</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0.6</td>
<td>0.3</td>
<td>0.3</td>
<td>0</td>
<td>0.4</td>
<td>0.3</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.6</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0.6</td>
<td>0.4</td>
<td>0.5</td>
<td>0.8</td>
<td>0.8</td>
<td>0.5</td>
<td>0.8</td>
<td>0.5</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0.9</td>
<td>1.0</td>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
<td>0.8</td>
<td></td>
</tr>
</tbody>
</table>

Table J.10: The proportion of runs for which the Dispersion Index of Polarization detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 50.
Table J.11: Simpson Index, N=50

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table J.12: The proportion of runs for which the Simpson Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 50.
Appendix X. *Detailed Simple model results*

### Table J.13: Kempenaer Index, N=50

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### Table J.14: The proportion of runs for which the Simpson Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 50.

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
**Table J.15:** Kappa coefficient, N=50

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0.1</td>
<td>0</td>
<td>0.1</td>
<td>0.3</td>
<td>0.2</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0.2</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0.3</td>
<td>0.2</td>
<td>0.3</td>
<td>0</td>
<td>0.4</td>
<td>0.3</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.6</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0.5</td>
<td>0.3</td>
<td>0.4</td>
<td>0.8</td>
<td>0.5</td>
<td>0.6</td>
<td>0.7</td>
<td>0.3</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0.9</td>
<td>1.0</td>
<td>0.6</td>
<td>0.9</td>
<td>1.0</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>0.8</td>
</tr>
</tbody>
</table>

**Table J.16:** The proportion of runs for which the Kappa coefficient detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 50.

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0.1</td>
<td>0</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.4</td>
<td>0.4</td>
<td>0.8</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.4</td>
<td>0.9</td>
<td>0.9</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0.2</td>
<td>0.4</td>
<td>0</td>
<td>0.4</td>
<td>0.5</td>
<td>0.6</td>
<td>0.9</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0.5</td>
<td>0.8</td>
<td>0.4</td>
<td>0.9</td>
<td>0.7</td>
<td>0.8</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0.9</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>0.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.6</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0.1</td>
<td>0.4</td>
<td>0.4</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0.3</td>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>
Table J.17: Dispersion Index of Polarization, N=100

<table>
<thead>
<tr>
<th>$w_e$ =</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0.1</td>
<td>0.1</td>
<td>0</td>
<td>0.2</td>
<td>0.4</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0.2</td>
<td>0.5</td>
<td>0.3</td>
<td>0.5</td>
<td>0.3</td>
<td>0.9</td>
<td>0</td>
<td>0.5</td>
<td>0.2</td>
<td>0.4</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0.6</td>
<td>0.7</td>
<td>0.7</td>
<td>0.9</td>
<td>0.7</td>
<td>0.8</td>
<td>0.7</td>
<td>0.8</td>
<td>0.8</td>
<td>0.9</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table J.18: The proportion of runs for which the DIP detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100.
### Table J.19: Simpson Index, N=100

<table>
<thead>
<tr>
<th>$w_e$ =</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### Table J.20: The proportion of runs for which the Simpson Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100.

<table>
<thead>
<tr>
<th>$w_e$ =</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0.4</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>
### Table J.21: Kempenaer Index, N=100

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### Table J.22: The proportion of runs for which the Kempenaer Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100.

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table J.22: The proportion of runs for which the Kempenaer Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100.
### Table J.23: Kappa coefficient, N=100

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.3</td>
<td>0.0</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0.2</td>
<td>0.5</td>
<td>0.4</td>
<td>0.4</td>
<td>0.1</td>
<td>0.9</td>
<td>0.5</td>
<td>0.2</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0.6</td>
<td>0.7</td>
<td>0.7</td>
<td>0.9</td>
<td>0.6</td>
<td>0.8</td>
<td>0.7</td>
<td>0.8</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

### Table J.24: The proportion of runs for which the Kappa coefficient detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100.

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.4</td>
<td>0.6</td>
<td>0.8</td>
<td>0.9</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0.2</td>
<td>0.4</td>
<td>0.6</td>
<td>0.9</td>
<td>0.8</td>
<td>1.0</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0.9</td>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0.1</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0.1</td>
<td>0.6</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>
Appendix K

Detailed Complex model results

Complex model results
**Table K.1:** Dispersion Index of polarization, complex model, N=10

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0.5</td>
<td>0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0.1</td>
<td>0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0.3</td>
<td>0</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
</tr>
</tbody>
</table>

**Table K.2:** The proportion of runs for which the Dispersion Index of Polarization detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100.

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0.3</td>
<td>0</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0.3</td>
<td>0.1</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
</tbody>
</table>
Table K.3: Simpson Index, complex model, N=10

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table K.4: The proportion of runs for which the Simpson Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100.
### Table K.5: Kempenaer Index, complex model, N=10

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### Table K.6: The proportion of runs for which the Kempenaer Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100.

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
### Table K.7: Kappa coefficient, complex model, N=10

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = 0.1$</td>
<td>0.1</td>
<td>0.2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>$p_0 = 0.2$</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = 0.3$</td>
<td>0.4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = 0.4$</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>$p_0 = 0.5$</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### Table K.8: The proportion of runs for which the Kappa coefficient detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100.

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = 0.1$</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0.3</td>
<td>0.1</td>
<td>0.3</td>
</tr>
<tr>
<td>$p_0 = 0.2$</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
<td>0.2</td>
<td>0</td>
<td>0.3</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>$p_0 = 0.3$</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
<td>0</td>
<td>0.3</td>
<td>0.4</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>$p_0 = 0.4$</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0.3</td>
<td>0.4</td>
<td>0.6</td>
<td>0.9</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = 0.5$</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.3</td>
<td>0.3</td>
<td>0.6</td>
<td>0.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = 0.1$</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0.2</td>
<td>0.4</td>
<td>0.7</td>
<td>0.8</td>
<td>0.8</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = 0.2$</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0.2</td>
<td>0.3</td>
<td>0.7</td>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = 0.3$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0</td>
<td>0.3</td>
<td>0.4</td>
<td>0.6</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = 0.4$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0</td>
<td>0.3</td>
<td>0.4</td>
<td>0.6</td>
<td>0.8</td>
<td>0.9</td>
</tr>
<tr>
<td>$p_0 = 0.5$</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.4</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table K.8: The proportion of runs for which the Kappa coefficient detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100.
### Table K.9: Dispersion Index of polarization, complex model, N=50

<table>
<thead>
<tr>
<th>$w_e$ =</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0.1</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0.3</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0.5</td>
<td>0.4</td>
<td>0.3</td>
<td>0.5</td>
<td>0.4</td>
<td>0.3</td>
<td>0.2</td>
<td>0.4</td>
<td>0.2</td>
<td>0.4</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0.7</td>
<td>0.7</td>
<td>0.5</td>
<td>0.9</td>
<td>0.4</td>
<td>0.3</td>
<td>0.5</td>
<td>0.8</td>
<td>0.9</td>
<td>0.8</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0.9</td>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
<td>0.8</td>
<td>1.0</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$w_e$ =</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.1</td>
<td>0.3</td>
<td>0.4</td>
<td>1.0</td>
<td>0.8</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
<td>0.3</td>
<td>0.3</td>
<td>0.7</td>
<td>0.4</td>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0.3</td>
<td>0.4</td>
<td>0.1</td>
<td>0.5</td>
<td>0.7</td>
<td>0.7</td>
<td>0.9</td>
<td>1.0</td>
<td>0.9</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0.7</td>
<td>0.5</td>
<td>0.7</td>
<td>0.6</td>
<td>0.8</td>
<td>0.7</td>
<td>0.9</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>1.0</td>
<td>0.9</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>0.9</td>
</tr>
</tbody>
</table>

### Table K.10: The proportion of runs for which the Dispersion Index of Polarization detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 50.
Appendix Y. Detailed Complex model results

Table K.11: Simpson Index, complex model, N=50

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table K.12: The proportion of runs for which the Simpson Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 50.

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.3</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.7</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0.6</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0.7</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>
**Table K.13:** Kempenaer Index, complex model, N=50

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table K.14:** The proportion of runs for which the Kempenaer Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 50.

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Appendix Y. Detailed Complex model results

Table K.15: Kappa coefficient, complex model, N=50

<table>
<thead>
<tr>
<th>$w_e$ =</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0.3</td>
<td>0.1</td>
<td>0.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0.5</td>
<td>0.2</td>
<td>0.2</td>
<td>0.5</td>
<td>0.4</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0.7</td>
<td>0.6</td>
<td>0.2</td>
<td>0.8</td>
<td>0.4</td>
<td>0.3</td>
<td>0.4</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0.9</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
<td>0.8</td>
<td>0.9</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table K.16: The proportion of runs for which the Kappa coefficient detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 50.

<table>
<thead>
<tr>
<th>$w_e$ =</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0.3</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>1.0</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0.3</td>
<td>0.2</td>
<td>0.3</td>
<td>0.2</td>
<td>0.6</td>
<td>0.4</td>
<td>0.7</td>
<td>0.9</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0.3</td>
<td>0.4</td>
<td>0.1</td>
<td>0.3</td>
<td>0.7</td>
<td>0.6</td>
<td>0.8</td>
<td>0.9</td>
<td>0.9</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0.7</td>
<td>0.4</td>
<td>0.7</td>
<td>0.5</td>
<td>0.7</td>
<td>0.7</td>
<td>0.9</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0.9</td>
<td>0.9</td>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
<td>0.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$w_e$ =</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0.4</td>
<td>0.5</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0.4</td>
<td>0.4</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0.5</td>
<td>0.5</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0.6</td>
<td>0.7</td>
<td>1.0</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>
Appendix Y. Detailed Complex model results

Table K.17: Dispersion Index of Polarization, complex model, N=100

<table>
<thead>
<tr>
<th>( w_e ) =</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_0 = .1 )</td>
<td>0.1</td>
<td>0.4</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.3</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>( p_0 = .2 )</td>
<td>0.5</td>
<td>0.6</td>
<td>0.5</td>
<td>0.3</td>
<td>0.6</td>
<td>0.4</td>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>( p_0 = .3 )</td>
<td>1.0</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
<td>0.8</td>
<td>0.9</td>
<td>0.6</td>
<td>0.8</td>
<td>0.8</td>
<td>1.0</td>
</tr>
<tr>
<td>( p_0 = .4 )</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>( p_0 = .5 )</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table K.18: The proportion of runs for which the Dispersion Index of Polarization detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100.

<table>
<thead>
<tr>
<th>( w_e ) =</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_0 = .1 )</td>
<td>0.1</td>
<td>0.4</td>
<td>0.3</td>
<td>0.4</td>
<td>0.9</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>( p_0 = .2 )</td>
<td>0.1</td>
<td>0.4</td>
<td>0.9</td>
<td>0.5</td>
<td>0.7</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>( p_0 = .3 )</td>
<td>0.7</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>( p_0 = .4 )</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>( p_0 = .5 )</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( w_e ) =</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_0 = .1 )</td>
<td>0.1</td>
<td>0.7</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>( p_0 = .2 )</td>
<td>0.3</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>( p_0 = .3 )</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>( p_0 = .4 )</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>( p_0 = .5 )</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>
### Table K.19: Simpson Index, complex model, N=100

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0.1</td>
</tr>
</tbody>
</table>

### Table K.20: The proportion of runs for which the Simpson Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100.

<table>
<thead>
<tr>
<th>$w_e$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.6</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.4</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0.3</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0.6</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table K.20: The proportion of runs for which the Simpson Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100.
### Table K.21: Kempenaer Index, complex model, N=100

<table>
<thead>
<tr>
<th>$w_e =$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### Table K.22: The proportion of runs for which the Kempenaer Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100.

<table>
<thead>
<tr>
<th>$w_e =$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table K.22: The proportion of runs for which the Kempenaer Index detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100.
Table K.23: Kappa coefficient, complex model, N=100

<table>
<thead>
<tr>
<th>$w_e$ =</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>0.5</td>
<td>0.6</td>
<td>0.7</td>
<td>0.8</td>
<td>0.9</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0.4</td>
<td>0.5</td>
<td>0.4</td>
<td>0.3</td>
<td>0.5</td>
<td>0.3</td>
<td>0.3</td>
<td>0.4</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>1.0</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
<td>0.7</td>
<td>0.8</td>
<td>0.6</td>
<td>0.7</td>
<td>0.8</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>0.9</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table K.24: The proportion of runs for which the Kappa coefficient detected significant synchrony in the uncoordinated condition (top), conformity bias condition (middle), and shared response condition (bottom) with population size 100.

<table>
<thead>
<tr>
<th>$w_e$ =</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0.1</td>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
<td>0.8</td>
<td>0.9</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0</td>
<td>0.3</td>
<td>0.8</td>
<td>0.5</td>
<td>0.5</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0.7</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$w_e$ =</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_0 = .1$</td>
<td>0</td>
<td>0.4</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .2$</td>
<td>0.3</td>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .3$</td>
<td>0.9</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .4$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$p_0 = .5$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>
Bibliography


[26] Shmuel Noah Eisenstadt. Sectarianism and heterodoxy in Jewish history: Some comparative

[27] Magnus Enquist, Pontus Strimling, Kimmo Eriksson, Kevin Laland, and Jonas Sjostrand.

[28] Kimmo Eriksson and Julie C Coultas. The advantage of multiple cultural parents in the

[29] Braudel Fernand. La méditerranée et le monde méditerranéen à l’époque de philippe ii,


[33] Russell D Gray and Quentin D Atkinson. Language-tree divergence times support the

[34] Russell D Gray, Alexei J Drummond, and Simon J Greenhill. Language phylogenies reveal


[39] Marvin Harris. 1968. the rise of anthropological theory.


